Peer Instruction: Ten years of experience and results

Catherine H. Crouch and Eric Mazur[a)]
Department of Physics, Harvard University, Cambridge, Massachusetts 02138

(Received 21 April 2000; accepted 15 March 2001)

We report data from ten years of teaching with Peer Instruction (PI) in the calculus- and algebra-based introductory physics courses for nonmajors; our results indicate increased student mastery of both conceptual reasoning and quantitative problem solving upon implementing PI. We also discuss ways we have improved our implementation of PI since introducing it in 1991. Most notably, we have replaced in-class reading quizzes with pre-class written responses to the reading, introduced a research-based mechanics textbook for portions of the course, and incorporated cooperative learning into the discussion sections as well as the lectures. These improvements are intended to help students learn more from pre-class reading and to increase student engagement in the discussion sections, and are accompanied by further increases in student understanding.
[DOI: 10.1119/1.1374249]

I. INTRODUCTION

In recent years, physicists and physics educators have realized that many students learn very little physics from traditional lectures. Several investigators have carefully documented college physics students' understanding of a variety of topics, and have concluded that traditionally taught courses do little to improve students' understanding of the central concepts of physics, even if the students successfully learn problem-solving algorithms.[1] Simultaneously, authors studying learning in higher education have established that students develop complex reasoning skills most effectively when actively engaged with the material they are studying, and have found that cooperative activities are an excellent way to engage students effectively.[2] In response to these findings, many pedagogies have been devised to improve student understanding of physics, ranging from modifications of traditionally taught courses to complete redesign of courses.[3]

Here we present the results of ten years of teaching the two introductory physics courses for nonmajors at Harvard University with one such method, Peer Instruction (PI). Peer Instruction modifies the traditional lecture format to include questions designed to engage students and uncover difficulties with the material.[4,5] Peer Instruction has also been used successfully at many other institutions and in upper-level courses; those results are described elsewhere.[6]

This paper is structured as follows. Peer Instruction is described in Sec. II. In Sec. III, we present data showing ongoing improvement of student understanding as we have refined both implementation and materials. We describe these refinements in detail in Sec. IV. Most notably, to help students learn more from pre-class reading, we have replaced reading quizzes with a modified form of the Warm-up exercises of the Just-in-Time-Teaching strategy[7] and we have used sections of a research-based mechanics text;[8] to increase student engagement in the discussion sections, we have incorporated the *Tutorials in Introductory Physics* (McDermott *et al.*[3]) and group problem-solving activities similar to those developed by Heller *et al.*[3] One of the strengths of PI is its adaptability to a wide range of contexts and instructor styles. In Sec. IV we also provide recommendations for such adaptation, and describe resources available for implementing PI.

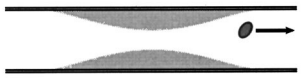
II. METHOD OVERVIEW

Peer Instruction engages students during class through activities that require each student to apply the core concepts being presented, and then to explain those concepts to their fellow students. Unlike the common practice of asking informal questions during a lecture, which typically engages only a few highly motivated students, the more structured questioning process of PI involves every student in the class. Although one of us (EM) developed PI for use in large lectures, many instructors have found it to be an effective approach for engaging students in small classes as well.[6]

A class taught with PI is divided into a series of short presentations, each focused on a central point and followed by a related conceptual question, called a ConcepTest (Fig. 1), which probes students' understanding of the ideas just presented. Students are given one or two minutes to formulate individual answers and report[9] their answers to the instructor. Students then discuss their answers with others sitting around them; the instructor urges students to try to convince each other of the correctness of their own answer by explaining the underlying reasoning. During the discussion, which typically lasts two to four minutes, the instructor moves around the room listening. Finally, the instructor calls an end to the discussion, polls students for their answers again (which may have changed based on the discussion), explains the answer, and moves on to the next topic. (A more detailed description of PI appears in Ref. 4.) Students are not graded on their answers to the ConcepTests, but do receive a small amount of credit for participating consistently over the semester. They also have a strong incentive to participate because the midterm and final exams include a significant number of ConcepTest-like questions.[10]

To free up class time for ConcepTests, and to prepare students better to apply the material during class, students are required to complete the reading on the topics to be covered *before class*. Learning from reading is a skill well worth developing, particularly because after college a great deal of ongoing learning takes place through reading. To help students identify and grasp the key points of the reading, as well as to provide an incentive for students to actually complete the reading, we give students credit for answering a few questions designed to help them think about the material. (This will be discussed further in Sec. IV A.)

A blood platelet drifts along with the flow of blood through an artery that is partially blocked by deposits.



As the platelet moves from the narrow region to the wider region, its speed

1. increases.
2. remains the same.
3. decreases.

Fig. 1. An example of a ConcepTest, taken from Ref. 4. (Answer: 3.)

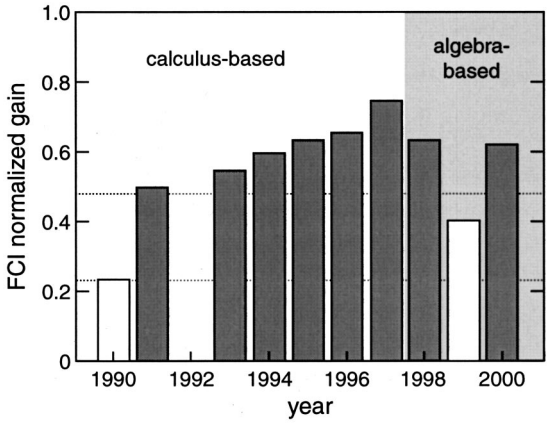


Fig. 2. Average Force Concept Inventory (Ref. 12) normalized gain $\langle g \rangle$ [Eq. (1)] for introductory calculus-based physics, Harvard University, Fall 1990–Fall 1997 (no data available for 1992), and for introductory algebra-based physics, Harvard University, Fall 1998–Fall 2000. Open bars indicate traditionally taught courses and filled bars indicate courses taught with PI. Dotted lines correspond to $\langle g \rangle = 0.23$, the typical gain for a traditionally taught course, and $\langle g \rangle = 0.48$, the typical gain for an interactive course (Hake–Ref. 1). The average pretest and posttest scores are provided in Table I.

III. RESULTS: IMPROVED STUDENT LEARNING

We find in both the algebra- and the calculus-based introductory physics courses[11] that our students' grasp of the course material improves according to a number of different measures: two standard tests, the Force Concept Inventory[12] and the Mechanics Baseline Test;[13] traditional examination questions; and ConcepTest performance, both during class and when tested for retention at the end of the semester. Although we see the most dramatic differences in student achievement between courses taught with traditional instruction and those taught with PI, we also observe continued improvement as we refine both pedagogy and ConcepTests.

We have improved our implementation of PI as follows: In 1993 and 1994, we refined the set of ConcepTests and the in-class questioning/discussion strategy. We began using a research-based text for one-dimensional mechanics in 1995.[8] In 1996, we introduced free-response reading assignments (described in Sec. IV A), and introduced cooperative learning into the discussion sections (Sec. IV B). Further improvement of the reading assignments took place in 1998. Because students learn from a wide range of activities in the course, it is plausible that student learning would continue to improve as more components of the course are modified to engage students more actively.

Over the seven years of results reported from the calculus-based course, five different instructors were involved, each using Peer Instruction with his or her own style; all but one of the instructors had extensive previous experience with traditional lecturing.[14] Thus the results reported here do not depend on a single particular instructor.

A. Conceptual mastery

Since 1990, we have given the Force Concept Inventory (FCI)[12] in our course at the beginning and at the end of the term. As shown in Table I, we find that the average pretest score $\langle S_{\text{pre}} \rangle$ (before instruction) for the calculus-based course stays essentially constant over the period tested (1990–1997).[15] Likewise, the difference between the average pretest scores for the algebra-based course in 1998 and 2000 is not statistically significant.[16]

The average posttest score $\langle S_{\text{post}} \rangle$ (after instruction) in the calculus-based course increases dramatically on changing from traditional instruction (1990) to PI (1991); as shown in Fig. 2 and Table I, the average normalized gain

$$\langle g \rangle = (\langle S_{\text{post}} \rangle - \langle S_{\text{pre}} \rangle)/(100\% - \langle S_{\text{pre}} \rangle) \tag{1}$$

doubles from 1990 to 1991, consistent with what has been observed at other institutions upon introducing interactive-engagement instruction (Hake—Ref. 1). With continued use of PI (1993–1997), along with additional improvements to the course, the normalized gain continues to rise. In 1998 and 2000 we see high normalized gains teaching the algebra-based course with PI, while the same course taught traditionally in 1999 by a different instructor produced a much lower, though still respectable, average normalized gain.

B. Quantitative problem solving

With PI, quantitative problem solving is de-emphasized in lecture; students learn these skills primarily through discussion sections and homework assignments. One way we assess our students' quantitative problem-solving skills is with the Mechanics Baseline Test (MBT).[13] Figure 3 and Table I show that the average score on this test in the calculus-based course increased from 66% in 1990 with traditional instruction to 72% in 1991 with the introduction of PI, and continued to rise in subsequent years, reaching 79% in 1997. Furthermore, student performance on the subset of MBT questions that require algebraic calculation also improved from 62% to 66% on changing from traditional lecturing to PI (also shown in Fig. 3 and Table I); for both traditional instruction and PI, the average score on those questions is about 5% lower than on the MBT overall.[17] In the algebra-based course taught with PI, the MBT scores are 68% in Fall 1998 and 66% in Fall 2000, consistent with Hake's findings that average scores on the MBT are typically about 15% lower than the FCI posttest score. The scores on the quantitative questions are 59% in Fall 1998 and 69% in Fall 2000. (No MBT data are available from the traditionally taught algebra-based course.)

For further comparison of conventional problem-solving skills with and without PI, in the calculus-based course, we administered the 1985 final examination, consisting entirely of quantitative problems, again in 1991 (the first year of instruction with PI). The mean score increased from 63% to 69%, a statistically significant increase (effect size 0.34),[18] and there are fewer extremely low scores. We also repeated

Table I. Force Concept Inventory (FCI) and Mechanics Baseline Test (MBT) results.[a]

Year	Method	FCI pre	FCI post	Absolute gain (post−pre)	Normalized gain $\langle g \rangle$	MBT	MBT quant. questions	N
Calculus-based								
1990	Traditional	(70%)	78%	8%	0.25	66%	62%	121
1991	PI	71%	85%	14%	0.49	72%	66%	177
1993	PI	70%	86%	16%	0.55	71%	68%	158
1994	PI	70%	88%	18%	0.59	76%	73%	216
1995	PI	67%	88%	21%	0.64	76%	71%	181
1996	PI	67%	89%	22%	0.68	74%	66%	153
1997	PI	67%	92%	25%	0.74	79%	73%	117
Algebra-based								
1998	PI	50%	83%	33%	0.65	68%	59%	246
1999	Traditional	(48%)	69%	21%	0.40	⋯	⋯	129
2000	PI	47%	80%	33%	0.63	66%	69%	126

[a]The FCI pretest was administered on the first day of class; in 1990 no pretest was given, so the average of the 1991–1994 pretest is listed. In 1995 the 30-question revised version was introduced (Ref. 15). In 1999 no pretest was given so the average of the 1998 and 2000 pretest is listed. The FCI posttest was administered after two months of instruction, except in 1998 and 1999, when it was administered the first week of the following semester to all students enrolled in the second-semester course (electricity and magnetism). The MBT was administered during the last week of the semester after all mechanics instruction had been completed. For years other than 1990 and 1999, scores are reported for matched samples for FCI pre- and posttest and MBT. No data are available for 1992 (EM was on sabbatical) and no MBT data are available for 1999.

individual problems from traditional exams on the midterms in the calculus-based course in 1991 (results reported in Ref. 4). Finally, in the second semester of the algebra-based course in Spring 2000 (electricity and magnetism), we included on the final exam one quantitative problem from the previous year, when a different instructor had taught the course traditionally. We found that the students taught with PI (Spring 2000, $N=155$) significantly outperformed the students taught traditionally (Spring 1999, $N=178$), averaging 7.4 out of 10 compared to 5.5 out of 10 (standard deviations 2.9 and 3.7, respectively). The improvement of the PI students over the traditional students corresponds to an effect size of 0.57. All measures indicate that our students' quantitative problem-solving skills are comparable to or better than those achieved with traditional instruction, consistent with the findings of Thacker *et al.*[19]

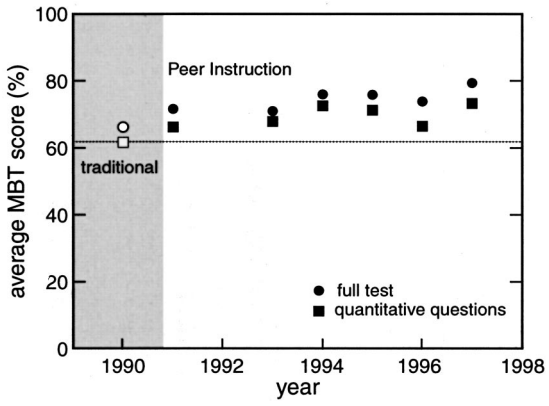


Fig. 3. Mechanics Baseline Test (Ref. 13) scores for introductory calculus-based physics, Harvard University, Fall 1990–Fall 1997. Average score on entire test (circles) and on quantitative questions (Ref. 17) only (squares) vs year are shown. Open symbols indicate traditionally taught courses and filled symbols indicate courses taught with PI. The dotted line indicates performance on quantitative questions with traditional pedagogy (1990).

C. ConcepTest performance

Students' responses to the ConcepTests themselves provide further insight into student learning. We analyzed student responses to all of the ConcepTests over an entire semester, and find that after discussion, the number of students who give the correct answer to a ConcepTest increases substantially, as long as the initial percentage of correct answers to a ConcepTest is between 35% and 70%. (We find that the improvement is largest when the initial percentage of correct answers is around 50%.[4]) In addition, the vast majority of students who revise their answers during discussion change from an incorrect answer to the correct answer. Figure 4 shows how students change their answers upon discussion for all of the ConcepTests used during the Fall 1997 semester. The answers are categorized as correct both before and after discussion (''correct twice''), incorrect before and correct after discussion (''incorrect to correct''), correct before and incorrect after discussion (''correct to incorrect''), or incorrect both before and after discussion (''incorrect twice''). Nearly half of the correct answers given were arrived at after discussion, and students changed from correct

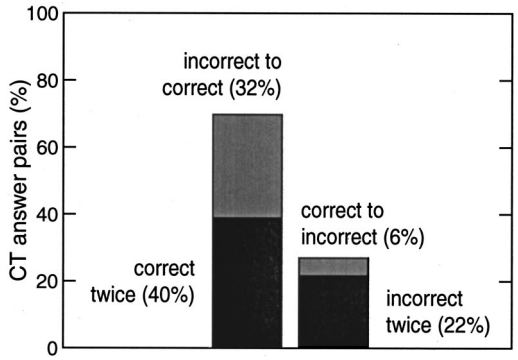


Fig. 4. Answers given to all ConcepTests discussed in Fall 1997, categorized as described in the text.

to incorrect answers during discussion only 6% of the time. We also examined the rate at which individual students give the correct answer prior to discussion,[5] and find that no student gave the correct answer to the ConcepTests prior to discussion more than 80% of the time, indicating that even the strongest students are challenged by the ConcepTests and learn from them.

In the algebra-based course, we examined student mastery of the ideas behind the ConcepTests by testing students at the end of the semester with free-response conceptual questions based on ConcepTests but with a new physical context. These questions thus required students to generalize the ideas they learned. We find that the number of students who successfully answer these questions (explaining their answer correctly as well as giving the correct answer) is comparable to the number who answer the ConcepTest correctly after discussion, and significantly greater than the number who answer the ConcepTest correctly before discussion, indicating that over the semester, students learn these ideas. (Of course, other elements of the course also help students learn these ideas; this study primarily indicates that students develop and retain real understanding of these concepts, which they lacked prior to discussion.) These results are presented in more detail elsewhere.[20]

IV. IMPLEMENTATION

As summarized in Sec. III, we have refined our implementation of Peer Instruction in three notable ways over the last several years. We have replaced reading quizzes with preclass Web-based assignments designed to help students think about the reading; we use a research-based mechanics text that is written to be read before class, rather than to serve primarily as a reference after lecture; and we have introduced cooperative activities in the discussion sections. Sections IV A and IV B elaborate on these improvements. Section IV C describes opportunities provided for learning quantitative problem-solving skills, and Sec. IV D describes strategies for motivating students.

Peer Instruction has been successfully adopted by hundreds of instructors at other institutions worldwide, and our communication with them indicates that one of the reasons for this widespread adoption is the ease of adapting PI to the local context.[6] An instructor can use ConcepTests developed elsewhere, write new questions, or use some of each. The choice of questions, the amount of time devoted to each question, the amount of lecturing, and the number of questions per class can and should be adapted to best suit a particular context and teaching style. Guidelines for such adaptations are given in Secs. IV E and IV F. For courses involving teaching assistants (TAs), strategies for TA training are given in Sec. IV G. Finally, Sec. IV H describes publicly available resources available for teaching with PI.

A. Reading incentives

In traditional introductory science courses, students generally read the textbook only after the lecturer has covered the topic (if ever). In a course taught with PI, students are expected to prepare for class by reading. This initial information transfer through reading allows the lectures to focus on the most important and difficult elements of the reading, perhaps from a different perspective or with new examples, and provide students with opportunities (in the form of ConcepTests) to think through and assimilate the ideas. To prepare themselves effectively for a PI class, students need both an incentive to complete the reading and guidelines for thinking about it before class.

Reading quizzes, which we used early on,[4] act as an incentive to complete the reading but do not help students think about it. In place of quizzes, in 1996 and 1997, we required students to write short summaries of what they read. We found, however, that most students did not write effective summaries.

The reading incentives we introduced in 1998, and have found most effective, are an adaptation of the Warmups from the Just-in-Time Teaching approach.[7] A three-question Web-based assignment is due before each class. All three questions are free response; the first two probe difficult aspects of the assigned reading, and the third asks, ''What did you find difficult or confusing about the reading? If nothing was difficult or confusing, tell us what you found most interesting. Please be as specific as possible.'' Students receive credit based on effort rather than correctness of their answers, which allows us to ask challenging questions, and vastly reduces the effort needed to grade the assignments.[21] Total credit for all of the reading assignments is worth 5% of the student's overall course grade (homework accounts for an additional 20% and exams for the remaining 75%).

Access to the students' responses to these questions allows the instructor to prepare for class more effectively; reading and thinking about students' questions gives the instructor insight into what students find difficult, complementing the instructor's ideas about what material needs most emphasis in class. Time spent preparing is comparable, because the instructor can spend less time reviewing other textbooks and notes for ideas on what should be covered, and this sort of preparation produces a class better suited to the students' identified needs. Student response to these reading assignments is particularly positive when their questions are answered (in class or by answers to FAQs posted on the course Web site).

B. Cooperative activities in discussion sections

Since 1996, to reinforce the interactive pedagogy of the lectures, we have structured discussion sections around cooperative activities as well. In the mechanics semester, students attend a weekly two-hour workshop (there is no separate laboratory period). Half of the workshop is devoted to conceptual reasoning and hands-on activities through the *Tutorials in Introductory Physics*[3] and half to quantitative problem solving. Cooperative problem-solving activities are described further in the next section.

C. Quantitative problem solving

As discussed in Sec. III, we find our students' problem-solving skills to be at least as good as before implementing PI. To achieve this, some direct instruction in quantitative problem-solving skills is necessary, and such instruction should help students connect qualitative to quantitative reasoning.[22] Students need opportunities to learn not only the ideas of physics but also the strategies employed by expert problem solvers; otherwise their main strategy often becomes finding a worked example similar to the problem at hand.

Two components of our course are designed to help students learn problem solving: discussion sections (''workshops'') and homework. The second half of the workshop

begins with the instructor solving a problem to illustrate the reasoning that goes into successful problem solving; the problem is chosen to be challenging without being tedious. Students spend the remainder of the hour working in groups on selected problems from the homework.[23] The instructor circulates around the classroom, asking students to explain their work and helping students through difficulties (by asking questions to lead them to the right answer, rather than by giving answers). At the end of the week, each student must turn in their own written solutions to the problems, and their homework solutions are graded individually on correctness.

The weekly homework assignments consist of ten problems, most of which are quantitative rather than conceptual. We provide the students at the beginning of the year with a handout on problem-solving strategies taken from Heller *et al.*[3] and encourage instructors to explicitly use the steps from the handout in solving the example problems. We also encourage students to attempt the homework before the workshop so that they can benefit most from group work.

D. Student motivation

It has been established[24] that students often require a period of adjustment to new methods of instruction before their learning improves. In the same fashion, when learning a new way to grip a tennis racquet, a tennis player is likely to play worse at first, and improve only after becoming comfortable with the new (and presumably better) grip. At such times, it is the coach's responsibility to encourage the player that this decline is a normal part of the learning process. Likewise, in the classroom, the instructor must not be discouraged by complaints such as, ''When are we going to do some *real* physics?'' and must continue to explain to students the reasons that the course is taught this way.[25]

Peer Instruction requires students to be significantly more actively involved and independent in learning than does a conventional lecture class. It is common for some or many students to be initially skeptical about this form of instruction.[26] Consequently, proper motivation of the students is essential. Motivation takes two forms: grading students on conceptual understanding, not just traditional problem solving, and setting the right tone in class from the start (including explaining the reasons for teaching this way). Including conceptual questions on exams makes it clear that the instructor is serious about the importance of conceptual understanding; providing equation sheets or making the exams open-book so that students do not need to memorize equations is also important. Giving an examination early in the semester is useful to communicate this from the start; distributing copies of past exams with the syllabus can also be helpful. Strategies for setting the right tone are given in *Peer Instruction: A User's Manual*.[4]

Student attitudes to a course taught with PI, as measured by student evaluations and by our interactions with students, have differed. In the calculus-based course, EM's average evaluation score—4.5 on a scale of 1–5[27]—did not change on introducing PI, and written comments on evaluations indicated that the majority of students appreciated the interactive approach of the course. For the algebra-based course, while still good, EM's average evaluation score dropped significantly, to 3.4;[28] although most students are satisfied with the course, there are more dissatisfied students than in the calculus-based course. Some of this dissatisfaction is not related to PI; the most frequent complaint about the algebra-based course is that it meets at 8:30 a.m. (the calculus-based course meets at 11 a.m.). We also surmise that students in the algebra-based course are on average less interested in the course and more intimidated by the material, since these students are primarily nonscience majors; the students in the calculus-based course are mostly honors biology or chemistry majors.

We also examined student attitudes by giving the concept and reality link clusters from the MPEX[29] to the algebra-based course in 1998. For both clusters, we found that the percentage of favorable responses remained exactly the same from the precourse to the postcourse survey (68% for concepts and 67% for reality link), and the percentage of unfavorable responses increased slightly (from 11% to 14% for concepts and from 12% to 15% for reality link; the remaining responses were neutral). Thus we find very little change in class attitudes over the semester. In their six-institution study, the MPEX authors found a small increase in favorable responses on the concept cluster and a small to moderate decrease in favorable responses on the reality link cluster.[29]

It is important to note that student evaluations and attitude are not a measure of student learning; as discussed in Sec. II, we saw high learning gains for the students in the algebra-based course in spite of lower perceived satisfaction overall. Other instructors report similar experiences.[30] Furthermore, research indicates that student evaluations are based heavily on instructor personality[31] rather than course effectiveness. We are nevertheless continuing to try to find strategies that will help motivate more of the students in the algebra-based course.

E. ConcepTest selection

Appropriate ConcepTests are essential for success. They should be designed to give students a chance to explore important concepts, rather than testing cleverness or memory, and to expose common difficulties with the material. For this reason, incorrect answer choices should be plausible, and, when possible, based on typical student misunderstandings. A good way to write questions is by looking at students' exam or homework solutions from previous years to identify common misunderstandings, or by examining the literature on student difficulties. ConcepTests should be challenging but not excessively difficult; as mentioned previously (Sec. III C and Ref. 4), 35%–70% of the students should answer correctly prior to discussion. If fewer than 35% of the students are initially correct, the ConcepTest may be ambiguous, or too few students may understand the relevant concepts to have a fruitful discussion (at least without some further guidance from the instructor). If more than 70% of the students can answer the question correctly alone, there is little benefit from discussion.

In a course with a large enrollment, it is often easiest for the instructor to poll for answers to multiple-choice questions. However, open-ended questions can also be posed using a variety of strategies. For example, the instructor can pose a question and ask students to write their answers in their notebooks. After giving students time to answer, the instructor lists several answer choices and asks students to select the choice that most closely corresponds to their own. Answer choices can be prepared ahead of time, or the instructor can identify common student answers by walking around the room while students are recording their answers and prepare a list in real time. This tactic works especially well when the answer is a diagram or graph.

It is possible to pose quantitative problems in a similar manner. Students need more than two minutes to work on such problems individually before discussion. One approach is to have students outline the strategy for solving a complex, multi-step problem; the instructor then shows a list of possible first steps and asks students which step to choose. (This can lead to interesting discussions, because for many problems, more than one strategy is possible.) The primary challenge in such problems should be to identify the underlying physics and develop a strategy for solving the problem. Equations should be readily available to the students either on the blackboard or in the textbook (if students bring their books to class).[32] If mathematical answer choices are provided, incorrect choices should be results obtained from making likely errors.

F. Time management

We typically devote one-third to one-half of class time to ConcepTests and spend the remainder lecturing. (The amount of time varies from class to class depending on the topic and the difficulty of the material.) Other instructors may use only one ConcepTest per class, or may spend nearly all class time on ConcepTests; regardless of the number, using ConcepTests leaves less time for traditional lecture presentation of material. The instructor therefore has two choices: (a) discuss in lecture only part of the material to be covered over the semester (and expect the students to learn the remainder from reading, problem sets, and discussion sections) or (b) reduce the number of topics covered during the semester. In the calculus-based course, we opted for the first strategy. In the algebra-based course, we followed the second, reducing the number of topics covered by 10%–15%[33] and covering those topics in more depth. The best approach depends on the abilities of the students and the goals of the course.

To make the most of class time, we streamline the lecturing component of class in several ways. Lectures include very few derivations; the instructor instead explains the strategy used to obtain a result from its starting point, highlighting the strategy and the conceptual significance. Students are expected to study derivations outside of class, when they can go at their own pace. If the derivation is not explained well in the text, the instructor provides a handout with more detailed comments. Because students are expected to read before class, less time is spent repeating definitions that are printed in the textbook. The instructor chooses quantitative examples for maximum physical insight and minimal algebra, and often works such examples in the process of explaining a related ConcepTest. Examples that are primarily mathematical can be presented in small discussion sections (where the instructor can tailor the presentation to the individual students present and answer their questions), or studied by students from the text or handouts.

G. Teaching assistant training

In courses involving teaching assistants (TAs), the TAs have a significant impact on students' experience. While many TAs are excited by the opportunity to engage their students more actively, some resist innovation and may communicate a negative attitude to the students. To avoid this problem as much as possible, it is vital to motivate TAs as well as students.[34] Before the course begins, we explain to our TAs the reasons for teaching with PI and give them the data on improved student learning. We also require our TAs to attend lecture, both so that they will be best able to help students and so that they see PI in action (which often convinces skeptical TAs).

One way to help TAs see the value of PI is to have them think about and discuss challenging ConcepTests, so that they experience the benefits of discussion. If such ConcepTests are related to the course material, this also makes them realize that they don't know everything already! (Questions on introductory fluid statics and dynamics are usually challenging for our TAs.) We hold a weekly meeting for our teaching staff, during which we go through the material to be covered the following week in section, emphasizing the pedagogy we wish them to use.

H. Resources

There are a number of resources available for implementing PI in introductory physics courses (as well as in chemistry and astronomy courses). *Peer Instruction: A User's Manual*[4] includes 243 ConcepTests developed for our introductory calculus-based physics for nonmajors, covering mechanics, electricity, magnetism, fluid statics and dynamics, oscillations and waves, geometrical and physical optics, and modern physics. A searchable database of ConcepTests on the Project Galileo Web site (http://galileo.harvard.edu; free registration required for access) includes over 800 physics ConcepTests, many developed at other institutions for either algebra- or calculus-based introductory physics, and some developed for nonintroductory courses. Utilities for this database allow the user to generate class-ready materials, such as pages for a course Web site, directly from the database. Links to separate databases of ConcepTests for astronomy and chemistry are also available. A resource Web site, http://galileo.harvard.edu/galileo/course/index.html, provides a full archive of our course materials, organized in the same manner as our course Web site.

V. CONCLUSIONS

We find that, upon first implementing Peer Instruction, our students' scores on the Force Concept Inventory and the Mechanics Baseline Test improved dramatically, and their performance on traditional quantitative problems improved as well. Subsequent improvements to our implementation, designed to help students learn more from pre-class reading and to increase student engagement in the discussion sections, are accompanied by further increases in student understanding. These results are not dependent on a particular instructor and are seen in both the algebra-based and calculus-based courses. Finally, with significant effort invested to motivate students, student reactions to PI are generally positive, though there are always some students resistant to being taught in a nontraditional manner, and we find more students are resistant in the algebra-based course than the calculus-based course.

ACKNOWLEDGMENTS

The authors would like to thank Professor Michael J. Aziz and other members of the Physics 1 and 11 teaching staff for ongoing partnership in developing Peer Instruction and ConcepTests; Emily Fair Oster and Cinthia Guzman for help

with data analysis; and Dr. Paul Callan, Adam Fagen, Professor Richard Hake, and Chris Schaffer for helpful discussions.

[a]Electronic mail: mazura@physics.harvard.edu

[1]For example, see I. Halloun and D. Hestenes, ''The initial knowledge state of college physics students,'' Am. J. Phys. **53** (11), 1043–1055 (1985); L. C. McDermott, ''Millikan Lecture 1990: What we teach and what is learned—Closing the gap,'' *ibid.* **59**, 301–315 (1991); R. R. Hake, ''Interactive-engagement vs. traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses,'' *ibid.* **66** (1), 64–74 (1998).

[2]D. W. Johnson, R. T. Johnson, and K. A. Smith, *Active Learning: Cooperation in the College Classroom* (Interaction Book Company, Edina, MN, 1991); R. T. Johnson and D. W. Johnson, ''Cooperative learning and the Achievement and Socialization Crises in Science and Mathematics Classrooms,'' from *Students and Science Learning: Papers from the 1987 National Forum for School Science* (AAAS, Washington, DC, 1987), and references therein.

[3]Examples include L. C. McDermott, P. S. Schaffer, and the University of Washington PERG, *Tutorials in Introductory Physics* (Prentice–Hall, Upper Saddle River, NJ, 1998); *Workshop Physics* (developed by P. W. Laws, R. Thornton, D. Sokoloff, and co-workers, and published by John Wiley); *Active Learning Problem Solving Sheets* (developed by A. van Heuvelen, Ohio State University); and numerous forms of Socratic dialogue, as in R. R. Hake, ''Socratic Pedagogy in the Introductory Physics Lab,'' Phys. Teach. **30**, 546–552 (1992), or group problem solving, as in Patricia Heller, Ronald Keith, and Scott Anderson, ''Teaching problem solving through cooperative grouping. Group versus individual problem solving,'' Am. J. Phys. **60** (7), 627–636 (1992), and Patricia Heller and Mark Hollabaugh, ''Teaching problem solving through cooperative grouping. 2. Designing problems and structuring groups,'' *ibid.* **60** (7), 637–644 (1992). Materials for these innovations are available by contacting the publishers or the developers; information on several innovations is also available at http://galileo.harvard.edu.

[4]Eric Mazur, *Peer Instruction: A User's Manual* (Prentice–Hall, Upper Saddle River, NJ, 1997). Additional information and resources for PI can be found at http://galileo.harvard.edu.

[5]Catherine H. Crouch, ''Peer Instruction: An Interactive Approach for Large Classes,'' Opt. Photonics News **9** (9), 37–41 (September 1998).

[6]Adam P. Fagen, Catherine H. Crouch, Tun-Kai Yang, and Eric Mazur, ''Factors That Make Peer Instruction Work: A 700-User Survey,'' talk given at the 2000 AAPT Winter Meeting, Kissimmee, FL, January 2000; and ''Peer Instruction: Results From a Range of Classrooms'' (unpublished).

[7]Gregor Novak, Evelyn Patterson, Andrew Gavrin, and Wolfgang Christian, *Just-in-Time Teaching: Blending Active Learning and Web Technology* (Prentice–Hall, Upper Saddle River, NJ, 1999), and http://webphysics.iupui.edu/jitt/jitt.html.

[8]Since 1995, we have replaced textbook readings on one-dimensional mechanics with a draft text written by Eric Mazur, in which concepts are introduced prior to the mathematical formalism, and many research findings of typical student difficulties are directly addressed in the text. In 1998 and 2000 this text was used for all topics in mechanics in the algebra-based course.

[9]Methods for polling for student answers include a show of hands or flashcards, classroom network systems, and scanning forms. A discussion of the pros and cons of each of these methods is given in Ref. 4; we used scanning forms combined with a show of hands in 1991 and classroom network systems thereafter. We did not see any significant changes in student learning on introducing the classroom network system, and find the main advantages of the network are anonymity of student responses and data collection; our experience indicates that the success of Peer Instruction does not depend on a particular feedback method.

[10]Exam questions are free-response and graded primarily on the quality of the student's explanation of the answer. In class, we typically use multiple-choice ConcepTests, for ease of polling students for their answers.

[11]The ''algebra-based'' course involves a very small amount of single-variable calculus, primarily derivatives and an occasional integral, in the second semester (electricity & magnetism). The students in this course have less facility with mathematical problem solving than in the calculus-based course.

[12]The FCI is a test of conceptual understanding of mechanics, written in ordinary language so that it can be given before as well as after mechanics instruction. The original version is published in D. Hestenes, M. Wells, and G. Swackhammer, ''Force Concept Inventory,'' Phys. Teach. **30** (3), 141–151 (1992). The test was revised in 1995 by I. Halloun, R. R. Hake, E. Mosca, and D. Hestenes; the revised version is printed in *Peer Instruction: A User's Manual* and can also be obtained from Professor Hestenes at Arizona State University. For nationwide data that have been gathered on student performance on the test, see Hake (Ref. 1). To maintain the validity of the tests, we do not use materials in class that duplicate FCI questions.

[13]D. Hestenes and M. Wells, ''A Mechanics Baseline Test,'' Phys. Teach. **30** (3), 159–166 (1992). This test is available from the same sources as the FCI (Ref. 12).

[14]In 1990, 1993, and 1994, the calculus-based course was co-taught by Eric Mazur and William Paul; in 1995, the course was taught by Eric Mazur; in 1991 and 1996, the course was co-taught by Michael J. Aziz and Eric Mazur; and in 1997, the year in which the highest FCI gains were obtained, the course was co-taught by Michael J. Aziz, Catherine H. Crouch, and Costas Papaliolios. Leadership of class periods was divided equally among co-instructors, with each instructor taking charge of the same number of classes. All instructors used Peer Instruction beginning in 1991.

[15]In 1994 we changed from the original (29-question) version of the FCI to the revised (30-question) version. An informal e-mail survey on the listserv PhysLrnR found that at institutions which have given the FCI for a number of years, instructors typically see both pretest and posttest scores drop by roughly 3% on changing to the revised version. We saw this drop in our pretest but not in our posttest scores. We thank Professor Laura McCullough of the University of Wisconsin-Stout for telling us about this survey.

[16]A *t*-test (two-tailed) was performed to determine the likelihood that the difference in average pretest scores is due to real differences between the populations of students rather than simply variation within the population of students. The p value was 0.26; a p value of 0.05 or less is generally agreed to indicate a statistically significant difference.

[17]The questions we identified as significantly quantitative are numbers 9, 11, 12, 17, 18, 23, 24, and 25 (eight in all).

[18]The exam distributions are published in Fig. 2.8 of Mazur (Ref. 4, p. 17). A *t*-test was performed to determine the likelihood that this increase in mean score was simply due to variation within the population of students rather than genuine improvement in understanding. The p value was 0.001, well below the threshold of 0.05 for statistical significance, indicating a statistically significant increase in mean score.

[19]B. Thacker, E. Kim, K. Trefz, and S. M. Lent, ''Comparing problem solving performance of physics students in inquiry-based and traditional introductory physics courses,'' Am. J. Phys. **62**, 627–633 (1994).

[20]Catherine H. Crouch, John Paul Callan, Nan Shen, and Eric Mazur, ''ConcepTests in Introductory Physics: What Do Students Get Out of Them?,'' American Association of Physics Teachers Winter 2000 Meeting, Kissimmee, FL, January 2000; ''Student Retention of ConceptTests'' (unpublished); for transparencies and preprints consult http://mazur-www.harvard.edu.

[21]To minimize grading work, the Web utility we have developed automatically assigns full credit to every completed answer, and a grader spot-checks answers via a Web interface, which takes relatively little time.

[22]Stephen Kanim, ''An investigation of student difficulties in qualitative and quantitative problem solving: Examples from electric circuits and electrostatics,'' Ph.D. thesis, University of Washington, 1999, and references therein.

[23]Guidelines for effective group work are found in Heller and Hollabaugh and Heller, Keith, and Anderson (Ref. 3), as well as Johnson, Johnson, and Smith (Ref. 2).

[24]Philip M. Sadler, ''Psychometric Models of Student Conceptions in Science: Reconciling Qualitative Studies and Distractor-Driven Assessment Instruments,'' J. Res. Sci. Teach. **35** (3), 265–296 (1998); ''How students respond to innovation,'' seminar at the 1998 NSF Faculty Enhancement Conference ''Teaching Physics, Conservation Laws First'' (audio available at http://galileo.harvard.edu/conference/program.html). The tennis instructor illustration is also courtesy of Professor Sadler (private communication).

[25]Richard M. Felder and Rebecca Brent, ''Navigating the Bumpy Road to Student-Centered Instruction,'' College Teaching **44**, 43–47 (1996).

[26]D. R. Woods, *Problem-Based Learning: How to Gain the Most from PBL* (self-published, 1994); R. J. Kloss, ''A nudge is best: Helping students

through the Perry scheme of intellectual development,'' College Teaching **42** (4), 151–158 (1994); Felder and Brent (Ref. 25).

[27]Students were asked to give their opinion of the statement ''The professor was an effective instructor overall'' on a five-point scale (1 = strongly disagree; 2 = disagree; 3 = neutral; 4 = agree; 5 = strongly agree). EM's average score in the calculus-based course for both traditional lecturing (one semester) and teaching with PI (six semesters) was 4.5, with standard deviations of 0.6 (traditional, $N = 125$) and 0.8 (PI, $N = 789$).

[28]Over three semesters in the algebra-based course (Fall 1998, Spring 2000, and Fall 2000; Spring 2000 was the electricity and magnetism semester of the course), which was taught only with PI, EM's average score was 3.5, standard deviation 1.2 ($N = 229$).

[29]Edward F. Redish, Jeffery M. Saul, and Richard N. Steinberg, ''Student Expectations in Introductory Physics,'' Am. J. Phys. **66** (3), 212–224 (1998).

[30]Linda R. Jones, J. Fred Watts, and Andrew G. Miller, ''Case Study of Peer Instruction in Introductory Physics Classes at the College of Charleston,'' Proceedings of Charleston Connections: Innovations in Higher Education, 2000 (submitted).

[31]Nalini Ambady and Robert Rosenthal, ''Half a Minute: Predicting Teacher Evaluations From Thin Slices of Nonverbal Behavior and Physical Attractiveness,'' J. Personality Soc. Psych. **64** (3), 431–441 (1993).

[32]Students do not necessarily remember equations in class, especially if they are not required to memorize equations. (Examinations in our course are open-book.)

[33]Lecture schedules for our courses are available online at http://galileo.harvard.edu/galileo/course/ in the ''Lectures'' area.

[34]Wendell Potter and collaborators at the University of California, Davis have developed an entire program of training teaching assistants in interactive teaching strategies, as reported at the AAPT Winter 2000 meeting.