Peer Instruction:
Engaging Students One-on-One, All At Once

Catherine H. Crouch [1], Jessica Watkins [2], Adam P. Fagen [3], and Eric Mazur [4**]

[1] Dept. of Physics & Astronomy, Swarthmore College, Swarthmore, PA 19081
[2] Div. of Engineering and Applied Sci., Harvard University, Cambridge, MA 02138
[3] Board on Life Sciences, National Research Council, Washington, DC 20001
[4] Dept. of Physics and Div. of Engineering and Applied Sciences,
Harvard University, Cambridge, MA 02138

Abstract:

Peer Instruction is an instructional strategy for engaging students during class through a structured questioning process that involves every student. Here we describe Peer Instruction (hereafter PI) and report data from more than ten years of teaching with PI in the calculus- and algebra-based introductory physics courses for non-majors at Harvard University, where this method was developed. Our results indicate increased student mastery of both conceptual reasoning and quantitative problem solving upon implementing PI. Gains in student understanding are greatest when the PI questioning strategy is accompanied by other strategies that increase student engagement, so that every element of the course serves to involve students actively. We also provide data on gains in student understanding and information about implementation obtained from a survey of almost four hundred instructors using PI at other institutions. We find that most of these instructors have had success using PI, and that their students understand basic mechanics concepts at the level characteristic of courses taught with interactive engagement methods. Finally, we provide a sample set of materials for teaching a class with PI, and provide information on the extensive resources available for teaching with PI.

* To whom correspondence should be addressed.
E-mail: mazur@physics.harvard.edu

Table of Contents

1. Peer Instruction: A widely used strategy for actively engaging students during lecture

In recent years, physicists and physics educators have realized that many students learn very little physics from traditional lectures. Several researchers 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 methods have been devised to improve student understanding of physics, ranging from modifications of traditionally taught courses to complete redesign of courses.[3]

Here we describe the method and results of the use one such pedagogy, Peer Instruction (PI), in both algebra- and calculus-based introductory physics courses. Peer Instruction modifies the traditional lecture format to include questions designed to engage students and uncover difficulties with the material.[4, 5, 6] We present evidence of its effectiveness through the results of more than ten years of teaching the two introductory physics courses for non-majors at Harvard University.[6] Additionally, we provide preliminary results of its use and effectiveness in introductory physics courses at other institutions.[7,8,9]

This paper is structured as follows. Peer Instruction is described in detail in Section 2. In Section 3, we present data from Harvard University showing ongoing improvement of student understanding as we have refined both implementation and materials.[6] We present evidence of the effectiveness of PI at other institutions in Section 4, including the results of a survey of 700 PI-users,[7,8,9] with data on students' conceptual understanding after being taught with PI at a variety of institutions. Section 5 provides recommendations for adaptation of PI in the classroom, and describes resources available for implementation.

2. How a course taught with Peer Instruction works

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 traditional lecture, which typically engages only a few highly motivated students, PI incorporates a more structured questioning process that involves every student in the class. A similar questioning process is also used with Thornton and Sokoloff's Interactive Lecture Demonstrations.[3] Although PI was developed at Harvard for use in large lectures, many instructors have found it to be an effective approach for engaging students in small classes as well, as will be discussed in Section 4.[9]

2.1. A Peer Instruction "Lecture" instructor's presentations are interspersed with questions for all students to answer

The goal of PI is to transform the lecture environment so that it actively engages students and focuses their attention on underlying concepts. Instead of presenting the level of detail covered in the textbook or lecture notes, lectures consist of a number of short presentations on key points, each followed by a ConcepTest – short conceptual questions, typically posed in a multiple-choice format, on the subject being discussed. (Figure 1) Therefore, each key point in a lecture takes roughly 15 minutes to cover: 7-10 minutes of lecturing, 5-8 minutes for a ConcepTest. One hour of lecturing can address about four key points.

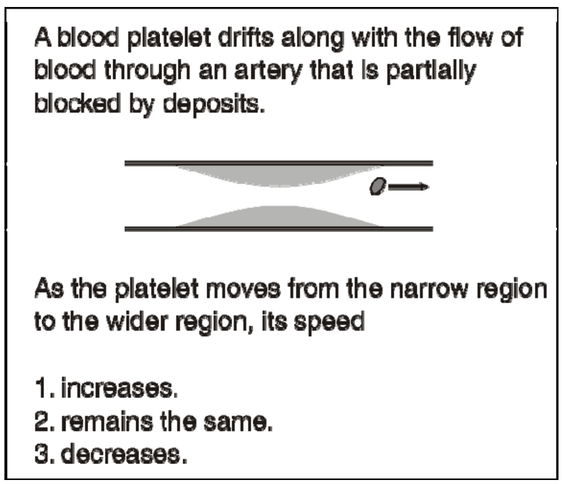


Fig. 1: An example of a ConcepTest, used by permission from ref. 6.

Each ConcepTest has the following general format:

1. Question posed 1 minute
2. Students given time to think 1-2 minutes
3. Students record/report individual answers

4. Neighboring students discuss their answers 2-4 minutes
5. Students record/report revised answers
6. Feedback to teacher: Tally of answers
7. Explanation of correct answer 2+ minutes

The students first consider the question on their own and are given one or two minutes to formulate individual answers and report 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 their own answer by explaining the underlying reasoning. Students are encouraged to "find someone who disagrees with you" for this discussion.

During the discussion, which typically lasts two to four minutes, the instructor moves around the room listening and, when necessary, asking questions to help the students in their thinking (if there are teaching assistants in the course, they do likewise). 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. Students are not graded on their answers to the ConcepTests, but do receive a small amount of credit for participating consistently over the semester. This method of questioning (a) forces the students to think through the arguments being developed, and (b) provides them (as well as the teacher) with a way to assess their understanding of the concept.

After this process, if most students choose the correct answer to the ConcepTest, the lecture proceeds to the next topic. If the percentage of correct answers after discussion is too low (perhaps less than 90%), the lecture slows down and goes into more detail on the same subject, and students' understanding is re-assessed with another ConcepTest. This repeat-when-necessary approach prevents a gulf from developing between the teacher's expectations and the students' understanding – a gulf that, once formed, only increases with time until the entire class is lost. A sample lecture plan, including ConcepTests, is included in section 5.

The balance between lecturing and questioning

Both at Harvard and at Swarthmore, 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.) There is a great deal of flexibility in the use of these conceptual questions; other instructors may use only one ConcepTest per class, or may spend nearly all class time on ConcepTests. However, regardless of the number, using ConcepTests leaves less time for traditional lecture presentation of material. The instructor therefore has two choices: (a) discuss only part of the material in lecture, 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 at Harvard, and more recently, in the sophomore electricity and magnetism course for physics majors at Swarthmore, we opt for the first strategy. In the algebra-based course at Harvard, we followed the second, reducing the number of topics covered by 10-15% and covering the remaining topics in more depth. The most suitable approach for a particular course depends on the abilities of the students and the goals of the course. A discussion of the topics covered in the Harvard calculus- and algebra-based introductory mechanics courses, along with syllabi for these courses, is provided in Appendices 1 and 2 to give two examples of what a course can cover when using Peer Instruction.

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. The instructor chooses quantitative examples for maximum physical insight and minimal algebra, and often solves such problems 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 in the section and answer their questions), or studied by students from the text or handouts.

How teaching changes with Peer Instruction

Peer Instruction lectures are much less rigid than those of the conventional method because, with the former, a certain amount of flexibility is necessary to respond to the sometimes unexpected results of the ConcepTests. The instructor usually needs to improvise more often than in conventional teaching. While this may seem a disturbing prospect at first, we have found that the added flexibility actually makes teaching *easier* in many ways as well as more challenging in others. During the periods of silence (when the students are thinking), the instructor has a break to breathe and think. During the convince-your-neighbors discussions, the instructor and teaching assistants can participate in some of the discussions. This participation is beneficial in two ways. First, it helps to hear students explain the answer in their own words. While the instructor’s explanations may be the most direct route from question to answer – the most efficient in terms of words and time – those provided by the students are often much more effective at convincing a fellow student, even if less direct. Sometimes the students offer a completely different perspective on the problem, which can help the instructor explain the concept better. In effect, the students can teach the teacher how to teach. What is also important is that by listening to students who have reasoned their way to the wrong answer, one can get a feel for what goes on in their minds. This involvement helps the instructor to better understand the problems the students are facing and address them directly in lecture. Finally, the personal interactions during the discussions can help the teacher keep in touch with the class.

Using ConcepTests elicits far more questions from the students than does traditional lecturing. Many of these questions are to the point and profound; the instructor should address as many of these as time permits. The increased engagement during class also prompts increased interaction among students and between students and instructor outside of class, through increased use of office hours, web-based discussion boards, e-mail, and so on.

2.2. ConcepTests: the cornerstone of teaching with Peer Instruction

What makes a good ConcepTest

Appropriate ConcepTests are essential for success. They should be designed to expose students' difficulties with the material, and to give students a chance to explore important concepts; they should not primarily test cleverness or memory. For this reason, incorrect answer choices must be plausible, and, when possible, based on typical student misunderstandings.

While there are no hard-and-fast rules for what makes a good ConcepTest, most satisfy a number of basic criteria:

- Focus on a single important concept, ideally corresponding to a common student difficulty
- Require thought, not just plugging numbers into equations
- Provide plausible incorrect answers
- Be unambiguously worded
- Be neither too easy nor too difficult

All these criteria directly affect feedback to the instructor. If more than one concept is involved in the question, it is difficulty for the instructor to interpret the results and correctly gauge understanding. If students can arrive at the answer by simply plugging numbers into equations, the answer does not necessarily reflect real understanding.

Extensive databases of ConcepTests for introductory physics, chemistry, and astronomy are publicly available, so faculty teaching such a course can simply make use of existing ConcepTests if they wish. Section 5 provides detailed information about these databases.

Writing good ConcepTests

For those writing their own, 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. Ideally, the incorrect answer choices should reflect students' most common misconceptions. One can attempt to formulate the incorrect answers (distracters) to each ConcepTest with this criterion in mind, but the ultimate source for distracters should be the students themselves. For instance, by posing an open-ended question on homework or an exam and then compiling the most frequent incorrect responses, a student-generated ConcepTest question that accurately mirrors common misconceptions is born.

Clarity is harder to gauge as a question is being created. Questions that may appear completely straightforward and unambiguous to the instructor are often misinterpreted by students. Needless to say, a question that is misinterpreted by students does not provide useful feedback; however, even ambiguous questions can provide a learning experience for students as they work to understand the confusing issues. In developing new ConcepTests, it is helpful when possible to "try out" the questions on a small group of students before class to make sure that there are no ambiguities in the wording and their interpretation is accurate.

ConcepTests should be challenging but not excessively difficult; we have observed that an initial correct response of roughly 35% to 70% prior to discussion seems to lead to the highest degree of engagement and the most effective discussions. 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. We discuss tailoring the ConcepTests to adequately challenge students further in Section 3, with data showing the pre- and post-responses for ConcepTests used over an entire semester.

At Swarthmore, PI has been used in a sophomore-level course for physics majors. For a course at this level, there has been some difficulty in developing ConcepTests that are suitable for students to answer in a minute or two but are difficult enough that only about half of the class gets them right before discussion. However, it has also been observed that even when most students give the correct answer on their own, they usually have to think hard to come up with the answer. Also, on course evaluations, students indicate that they found the ConcepTests useful in understanding the material, so the ConcepTests still appear to be beneficial to the students. If most of the class gets the question right, the discussion stage is skipped, but a student who gave the correct answer is asked to explain the answer to the class.

Beyond multiple-choice: Alternative ConcepTest formats

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 may 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) as students do not necessarily remember equations in class — especially if students are allowed to use a reference list of equations during exams. If mathematical answer choices are provided, incorrect choices should include results obtained from making errors in reasoning rather than arithmetical errors. It is also possible to simply not provide answer choices, but let students work on the problem for a while and then discuss their progress with their neighbors.

2.3. Use of demonstrations with Peer Instruction

Lecture demonstrations can be used effectively in combination with ConcepTests, with one leading into the other. For instance, a demonstration can be used to lead into a question that forces students to think about what they have just observed. Working the other way, students can be asked to think about a particular question and a demonstration can be used to answer it, in a manner similar to the Interactive Lecture Demonstrations.[3]

Crouch, Fagen, Callan and Mazur recently examined how the mode of presenting classroom demonstrations can affect student learning. These results are presented in detail elsewhere;[10] here we summarize the most important findings. In the discussion sections, demonstrations were presented in one of three different modes: *observe, predict,* and *discuss.* (Some sections were not shown the demonstration at all to serve as a control.) All students observed the demonstration and heard the instructor's explanation of the outcome. The *predict* group were also given a couple of minutes before the demonstration to predict the outcome and record their predictions. Students in the *discuss* group predicted the outcome, and after the demonstration, discussed the outcome with fellow students before hearing the instructor's explanation.

At the end of the semester, all students were tested on their ability to predict the outcomes of the same demonstrations shown in class and to explain their predictions. Although all students who saw the demonstration could predict the outcomes more successfully than the students who had not seen the demonstration, students in the *observe* group, who saw the demonstration without opportunity to predict the outcome or discuss the results, performed only marginally better in explaining the reason for their predictions than those who had not seen the demonstrations at all, and the difference was not statistically significant. In sharp contrast, the *predict* and *discuss* groups were able to give correct explanations at significantly higher rates than the *observe* group or the group who did not see the demonstration at all. On average the *predict* group required only an additional 2 minutes and the *discuss* group an additional 8 minutes over the *observe* mode; the improvement in student learning therefore seems worth the short amount of time needed to administer a quick ConcepTest to students asking them to predict the outcome.[10]

Not only does prefacing demonstrations with ConcepTests improve student understanding of the demonstrations, generally student interest in the demonstrations increases as well. When we began teaching with Peer Instruction, we observed that demonstrations which had previously elicited little student enthusiasm were now accompanied by much more excitement. For example, in teaching DC circuit analysis conventionally, demonstrating that light bulbs in parallel burn more brightly than light bulbs in series drew little, if any, student response. However, when the demonstration is preceded by a ConcepTest that asks them to predict the outcome, the excitement of the students on discovering that their answer was correct is audible!

2.4. Methods for polling students

One of the great advantages of Peer Instruction is that the ConcepTest answers give the instructor immediate feedback on student understanding. Tallying the answers can be accomplished in a variety of ways depending on setting and purpose:

1) *Show of hands.*

A show of hands after students have answered a question for the second time is the simplest feedback method. It gives the instructor a feel for the level of student understanding of the topic and allows the instructor to pace the lecture accordingly. The main drawbacks are the lack of accuracy and the discomfort that some students may feel responding in front of their peers. Additionally, a show of hands before the convince-your-neighbors discussion can influence the discussion and subsequent polling.

2) *Flashcards.*

When using flashcards for feedback, each student is provided with a set of numbered cards, and when the instructor asks students to report their answers, all students simultaneously show the flashcard with the number of their answer.[11] Flashcards thus address two of the main problems with using a show of hands; because all cards are raised at once, it is relatively straightforward for the instructor to estimate the fraction of students with the correct answer, and because students cannot easily read each other's flashcards, it is difficult for the students to figure out[11] which answer is the most popular. This main shortcoming of this method is the lack of a permanent record (unless the class is small enough to record individual answers.) One user of Peer Instruction suggested in response to our survey[9] that a digital picture be taken to record student responses both before and after discussion.

3) *Scanning forms.*

If an accurate record of student responses is desired, students can report their answers on scanning forms. Both before and after discussion, students mark their answers to the ConcepTests on these forms. (Students can also be asked to record their degree of confidence in their answers, providing additional data.) This method yields an enormous body of data on attendance, understanding, improvement, and the short-term effectiveness of the *Peer Instruction* periods. The drawbacks are that it requires some work after each lecture and that there is a delay in feedback, the data being available only after the forms are scanned. Additionally, students may not commit to an answer, as they do not have to report their response immediately. For these reasons, when using scanning forms, it is helpful to ask for a show of hands or flashcards as well.

4) *Classroom networks.*

Technology has afforded instructors convenient ways to poll answers and receive immediate feedback, through the Personal Response System (PRS)[12] and other similar technologies. The PRS system allows students to enter their answers to the ConcepTests, as well as their level of confidence in their answers, through infrared or radio-frequency wireless transmitters, often called "clickers." Additionally, if wireless Internet access is available, students can use their cell phones, PDAs, or laptop computers to respond to ConcepTests via a web site, as described in Section 5. Student responses are relayed to the instructor on a computer screen and can be projected for the students to see after discussion. The main advantage of these systems is that accurate results are instantly available to the instructor. The system displays a count and a histogram of student responses to the instructor as the responses arrive, so the instructor knows when most students have answered, and can decide whether to proceed with discussion based on the number of students giving correct answers. However, the students cannot see the histogram, so their discussions are not influenced by knowing which answer was most commonly given by their classmates. In addition, student information (such as their name and class participation history) is available to the instructor, making large classes more personal. The classroom networking system also facilitates gathering data for research purposes

As discussed further in Section 4, PI has been used successfully with each of these feedback methods; technology is not required for successful implementation.[9]

2.5. Changing student study habits: Getting students to read before class

Pre-class reading: some use it, some don't

To free up class time for ConcepTests, and to prepare students better to apply the material during class, at Harvard, students are required to complete the assigned reading on the topics to be covered before class. Learning from reading is a skill well worth developing, because after college, a great deal of ongoing learning takes place through reading. In traditional introductory science courses, students generally read the textbook only after the lecturer has covered the topic (if ever). If students read effectively before class, less class time can be spent introducing definitions and basic concepts and equations that are easily accessible in the textbook. Lectures can then 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.

Unfortunately, first- and second-year college students have typically not learned how to read science textbooks effectively, and they are unlikely to read without an incentive to do so. (Many introductory physics textbooks are also more encyclopedic than readable.) Consequently, at Harvard we provide a set of incentives and guidance for pre-class reading modeled on the Just-in-Time Teaching approach of Patterson, Novak, Gavrin, and Christian;[13] CHC has also used this approach at Swarthmore in the introductory physics course for non-majors. However, course evaluations at both Swarthmore and Harvard indicate that students are divided on whether they find pre-class reading helpful. Some feel that they don't understand the reading until after class, even with guidance. It would not be surprising if some students are less able to benefit from pre-class reading than others due to differences in learning style or physics background. In addition, many instructors find that Peer Instruction works effectively without expecting students to read before class.[8,9] Consequently, it should be clear that assigning pre-class reading, although helpful, is not required for using Peer Instruction.

Incentives and guidance for reading before class

Reading quizzes, which we used early on at Harvard,[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.[13] 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 asking challenging questions, and vastly reduces the effort needed to grade the assignments.[14] Total credit for all of the reading assignments is worth 10% of the student's overall course grade (homework accounts for an additional 20% and exams, laboratories, and classroom participation for the remaining 70%).

Access to the students' responses to these questions allows the instructor to prepare for class more effectively by providing insight into what students find difficult, complementing the instructor's ideas about what material needs most emphasis in class. Time spent preparing is comparable to that required for a traditional lecture class; the instructor spends less time reviewing other textbooks and notes for ideas on what should be covered, and more time finding out from the students what they understand. 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).

Web interfaces for reading assignments

Many web-based courseware systems exist that can be used to administer these assignments. The Interactive Learning Toolkit (ILT) is a web-based classroom management system that was developed specifically to facilitate implementing Peer Instruction and Just-in-Time-Teaching[13] as well as managing classroom response systems; it provides an interface for administering these reading follow-up assignments, as well as many other tools designed to save the instructor time and effort and to facilitate interaction between the students and the instructor (and among students as well). More information about the ILT can be found in Section 5.

2.6. Interactive approaches to discussion sections

At Harvard, beginning in 1996, we have organized the discussion sections around cooperative activities to reinforce the interactive pedagogy of the lectures. 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 in the next section.

2.7. Teaching quantitative problem solving as part of a Peer Instruction course

As will be discussed in section 3, we find our students' problem-solving skills to be at least as good as they were in the same course taught by traditional lecturing. To achieve this, some direct instruction in quantitative problem-solving skills is necessary, and such instruction should help students connect qualitative to quantitative reasoning.[15] 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 the Harvard 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.[16] 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, students must turn in their own written solutions to the problems, and their homework solutions are graded individually on correctness.

The weekly homework assignments include ten quantitative problems. 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 the teaching assistants leading the workshops 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.

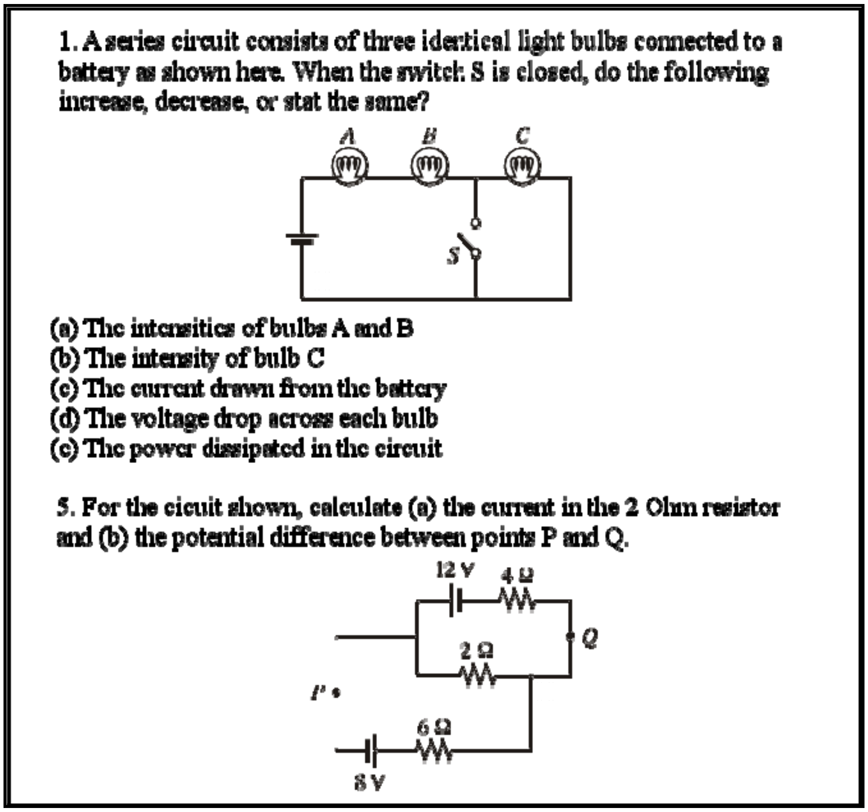
2.8. Determining student grades: exams include both conceptual questions and quantitative problems

When teaching with PI, examinations should include both conceptual short essay and quantitative problem-solving questions. This mix is essential because exams drive the way in which many students study. Examination questions are therefore a good way to make students realize the importance of conceptual understanding. At the beginning of the term, we distribute copies of past exams (with solutions) to the students and call attention to the conceptual problems.

Faculty often expect that conceptual questions will make examinations easier, but the opposite is true for those students whose approach to physics consists of searching for the right equation and then plugging numbers in. Only those who understand the underlying physics consider conceptual questions straightforward. Even with our conceptual emphasis during lecture, we find that students perform significantly better on the quantitative questions, as observed on the Harvard final exam in 2003.[17]

In the spring of 1991 at Harvard, while teaching the second semester of the calculus-based course traditionally, simple qualitative questions were paired with more difficult quantitative problems on the same physical concept. An example of such a pair of questions is provided in Figure 2. These questions were given as the first and last problem on one of the midterms (the other three problems on the examination, which were placed between these two, dealt with different subjects). Question 1 is purely conceptual and requires only a knowledge of simple circuits. Question 5 asks the students to perform a conventional calculation that involves the same underlying ideas; it requires setting up and solving two equations using Kirchoff's laws. Most physicists would consider question 1 easy and question 5 more difficult. As the results in Figure 3 indicate, however, students in a conventionally taught class would disagree.

Analysis of the responses reveals the reason for the large peak at 2 for the conceptual question: Over 40% of the students believed that closing the switch doesn't change the current through the battery but that the current splits into two at the top junction and rejoins at the bottom! In spite of this misconception, many still managed to correctly solve the mathematical problem.


1. A series circuit consists of three identical light bulbs connected to a battery as shown here. When the switch S is closed, do the following increase, decrease, or stat the same?

(a) The intensities of bulbs A and B
(b) The intensity of bulb C
(c) The current drawn from the battery
(d) The voltage drop across each bulb
(e) The power dissipated in the circuit

5. For the cicuit shown, calculate (a) the current in the 2 Ohm resistor and (b) the potential difference between points P and Q.

Fig. 2: Conceptual (top) and conventional question (bottom) on the subject of dc circuits. These questions were given on a written examination in 1991. Used by permission from ref. 4.

A proper balance between computational and conceptual problems is important. For our calculus-based introductory physics course, we administered two midterm examinations and one final examination. On a midterm, four or five out of seven questions are conceptual; on the final six out of twelve. Each problem carries the same weight because giving conceptual problems less weight favors those who manage to solve problems without understanding what they are doing.

In the electricity and magnetism course for sophomore physics majors at Swarthmore, it has been observed that especially on the first exam of the semester, some students perform very poorly on the conceptual examination questions because they do not write clear answers, even though these same students can give reasonably good explanations of their thinking in person. One reason this problem may not have been observed at Harvard is that there, students gain practice writing explanations on both the homework and the Tutorials in Introductory Physics, including conceptual homework associated with the Tutorials, while in the sophomore majors' course at Swarthmore, the Tutorials are not used because they do not match the level of the course, and thus the students' only practice writing explanations is as part of homework problems. (It is also possible that this same problem exists at Harvard, but is less easily identified because the large class size means that the instructor knows fewer of the students well.) This indicates that students need to have regular assignments that require them to write explanations of their ideas; otherwise they are being tested on something that they have not practiced adequately.

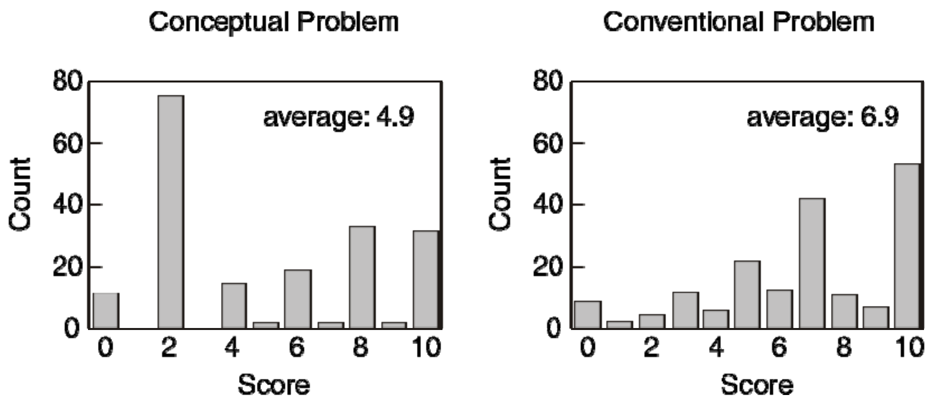


Fig. 3: Test scores for the problems shown in Figure 2. For the conceptual problem, each part was worth a maximum of 2 points. Used by permission from ref. 4.

2.9. Student attitudes in a Peer Instruction course

Motivating students thoroughly and from the start

Philip Sadler has established[18] 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.[19]

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.[20] 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. As mentioned in the previous section, we distribute copies of past exams on the first day of class so that students will see that understanding concepts will be required on the exams. One can also give an examination early in the semester.

It is important to set the tone starting on the first day of class. Students need to be told that they will not be lectured to straight out of the instructor’s notes or out of the textbook. As described in the *Peer Instruction User’s Manual*:[4]

> “I argue that it would be a waste of their time to have me simply repeat what is printed in the textbook or the notes. To do so implies that they are unable to read, and they ought to be offended when an instructor treats them that way. I explain how little one tends to learn in a passive lecture, emphasizing that it is not possible for an instructor just to pour knowledge in their minds, that no matter how good the instructor, *they* still have to do the work. I challenge them to become critical thinkers, explaining the difference between plugging numbers into equations and being able to analyze an unfamiliar situation.”

Creating a cooperative atmosphere

It is equally important to create a cooperative atmosphere in the classroom. Introductory science courses commonly have the reputation of being extremely competitive. This is detrimental to *Peer Instruction* as competition is incompatible with collaboration. One way to defuse such an atmosphere in a large class is with an absolute grading scale. After analyzing grades from the Harvard calculus-based course over several years, we found that averages tend to fluctuate very little from year to year. At

Harvard, we therefore issue worksheets that allow students to track their progress and determine their final grade on an absolute scale. (In practice, we make this scale slightly tougher than the one actually used to determine final grades, to give us a small amount of flexibility to be lenient.) When students know that no one's grade will go down because others have done better, the competition is removed from the classroom atmosphere.

The same atmosphere of cooperation is required for the convince-your-neighbors discussions; therefore students should not be graded on the correctness of their answers on ConcepTests. Instead, we give bonus points for participation, as we have found that providing an incentive for students to participate is beneficial.[21] For students who do not participate in the ConcepTests, their exams determine a somewhat higher fraction of their grade. Most students are eager to gain the edge on a good grade provided by simply participating in the ConcepTests.

Data on student attitudes from the Harvard courses

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[22]—4.5 on a scale of 1 to 5—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;[23] 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 non-science majors; the students in the calculus-based course are mostly honors biology or chemistry majors.

We also examined student attitudes by surveying students in the algebra-based course in 1998 using the concept and reality link clusters from the MPEX.[24] For both clusters, we found that the percentage of favorable responses remained exactly the same from the pre-course to the post-course 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.[24]

It is important to note that student evaluations and attitude are not a measure of student learning; as discussed in Section 3, 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.[25] Furthermore, research indicates that student evaluations are based heavily on instructor personality[26] 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.

2.10. Training teaching assistants to support Peer Instruction effectively

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.[27] 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.) At Harvard, 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.

3. Data from Harvard University: improved student learning

We find in both the algebra- and the calculus-based introductory physics courses[28] that our students' grasp of the course material improves according to a number of different measures: two standard tests, the Force Concept Inventory[29] and the Mechanics Baseline Test;[30] 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.[31] In 1996, we introduced free-response reading assignments (described in section 2), and introduced cooperative learning into the discussion sections (also described in section 2). 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.[32] Thus the results reported here do not depend on a single particular instructor.

3.1. Student mastery of concepts as measured by the Force Concept Inventory

Since 1990, we have given the Force Concept Inventory (FCI)[29] in our course at the beginning and at the end of the term. As shown in Table 1, we find that the average pretest score $\langle S_{pre} \rangle$ (before instruction) for the calculus-based course stays essentially constant over the period tested (1990 to 1997).[33] Likewise, the difference between the average pretest scores for the algebra-based course in 1998 and 2000 is not statistically significant.[34]

The average posttest score $\langle S_{post} \rangle$ (after instruction) in the calculus-based course increases dramatically on changing from traditional instruction (1990) to PI (1991); as shown in Figure 4 and Table 1, the average normalized gain

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

doubles from 1990 to 1991, consistent with what has been observed at other institutions upon introducing interactive-engagement instruction.[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.

3.2. Student mastery of quantitative problem solving: examinations and the Mechanics Baseline Test

With PI, quantitative problem solving is de-emphasized in lecture; students learn these skills primarily through discussion sections and homework assignments. To compare 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),[35] and there were fewer extremely low scores. We also repeated individual problems from traditional exams on the midterms in the calculus-based course in 1991, and again found that students taught with PI achieved comparable or better scores (results reported in ref. 4).

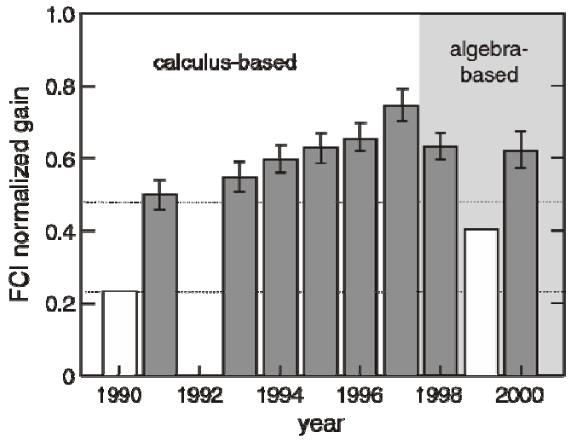


Fig. 4: Average Force Concept Inventory [ref. 29] normalized gain <*g*> [Eq. (1)] for introductory calculus-based physics, Harvard University, Fall 1990–1997 (no data available for 1992), and for introductory algebra-based physics, Harvard University, Fall 1998–Fall 2000. Open bars indicate traditionally taught courses and filled bars indicate courses taught with PI. The H-shaped bars indicate the 95% confidence interval around the normalized gain. Dotted lines correspond to <*g*>=0.23, the typical gain for an interactive course [Hake-Ref. 1]. The average pretest and posttest scores are provided in Table 1. Used by permission from ref. 6.

In addition, 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.[6]

Table 1: Force Concept Inventory (FCI) and Mechanics Baseline (MBT) results from the Harvard University. 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. 29]. 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 were obtained in 1992 (EM was on sabbatical) and no MBT data were obtained in 1999. Used by permission from ref. 6.

Year	***Method***	***FCI pre***	***FCI post***	***Absolute gain (post-pre)***	***Normalized gain <g>***	***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%	0.50	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

Another way we assess our students' quantitative problem-solving skills is with the Mechanics Baseline Test (MBT).[30] Figure 5 and Table 1 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. 5 and Table 1); for both traditional instruction and PI, the average score on those questions is about 5% lower than on the MBT overall.[36] 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.)

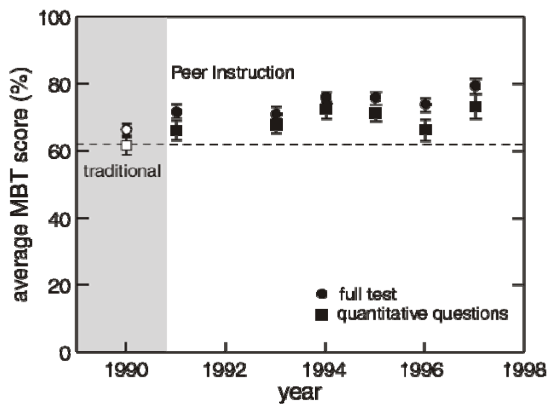


Fig. 5: Mechanics Baseline Test [ref. 30] scores for introductory calculus-based physics, Harvard University, Fall 1990–Fall 1997. Average score on entire test (circles) and on quantitative questions [ref. 36] only (squares) vs. year are shown. Open symbols indicate traditionally taught courses and filled symbols indicate courses taught with PI. The H-shaped bars indicate the 95% confidence intervals around the mean scores. The dotted line indicates performance on quantitative questions with traditional pedagogy (1990). Used by permission from ref. 6.

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.*[37]

3.3. Student learning of concepts: responses to ConcepTests

Students' responses to the ConcepTests 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%, as introduced in Section 2. (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 6 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 change from correct 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 gives 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.

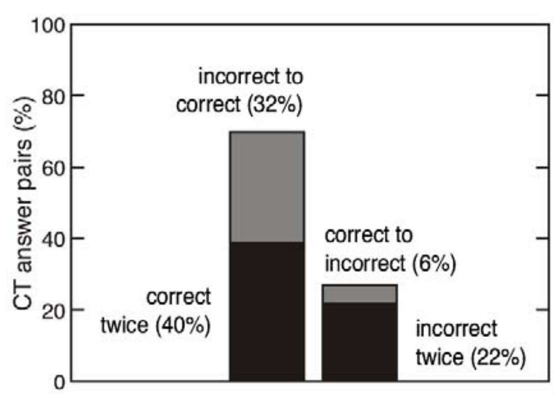


Fig. 6: Answers given to all ConcepTests discussed in Fall 1997, categorized as described in the text. Used by permission from ref. 6.

In the algebra-based course at Harvard, Crouch *et al.*[38] examined student mastery of the ideas behind the ConcepTests by assessing students' understanding at the end of the semester with free-response conceptual questions based on the ConcepTests but with a new physical context. These questions require students to generalize the ideas they learned to a different physical situation. An example of such a paired ConcepTest and exam question is provided in Figure 7.

ConcepTest:

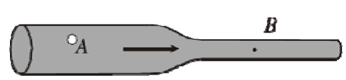
A blood platelet drifts along with the flow of blood through an artery that is partially blocked by deposits.

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

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

Exam Question:

Consider a small bubble of air trapped inside a pipe of non-uniform cross-section as shown in the figure, filled with flowing water. You may ignore the viscosity of the water.



The bubble travels along with the water around it; it expands in regions of low pressure and shrinks in regions of high pressure. As the bubble moves from point A to point B, how does its size change, if at all? Explain your answer briefly.

Fig. 7: ConcepTest (top) and examination question (bottom) on fluid dynamics, used in the study of ConcepTest retention.

Student responses to the exam questions and ConcepTests were then analyzed to compare the rates of correct answers to the ConcepTest before and after discussion to the rates of correct answers on the exam questions. The results are provided in Table 2. Only responses from students who both answered the ConcepTest and answered the exam question were analyzed. The exam questions were graded on a scale of 0 to 5, with 5 indicating completely correct and 4 indicating only minor errors that did not reflect a lack of understanding of a significant point; thus answers receiving scores of 4 or 5 were classified as correct. As the exam questions required students to explain their reasoning, only answers that gave a correct explanation were considered correct.

Table 2: Performance on ConcepTests and corresponding examination questions.
All percentages represent the percentage of students who answered correctly. The first three lines are questions asked on midterms and the later four lines are questions asked on the final exam. Results first reported in ref. 38.

Topic	***pre discussion***	***post discussion***	***on exam***
average velocity	54%	89%	77%
v(t) graph	34%	56%	68%
free-body diagram	19%	32%	46%
bubble	26%	59%	60%
wave on string	63%	95%	83%
seesaw	9%	46%	32%
oscillator	26%	53%	50%
total	34%	63%	61%

The results show that the number of students who answered the exam questions correctly is comparable to the number who answered the corresponding ConcepTest correctly after discussion, and significantly greater than the number who answered the ConcepTest correctly before discussion, indicating that over the semester, students gain understanding of 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. This observation is true over a wide range of levels of difficulty of the question.

3.4. The effect of Peer Instruction on the gender gap

Research on retention of female students in the physical sciences suggests that interactive teaching methods, a non-competitive atmosphere, and a conceptual emphasis should all make a course more welcoming to female students.[39] In the calculus-based physics course for non-majors at Harvard University, we indeed found that while male students outperformed female students on the Force Concept Inventory posttest when the course was taught traditionally, male and female students performed *equally* well when the course was taught with Peer Instruction in lecture and a mix of interactive strategies during the discussion sections. Furthermore, in the traditionally taught course there were many female students who scored below 60% on the Force Concept Inventory posttest and relatively few who scored above 85%; in the interactive course, there were no female students (and only a couple of male students) who scored below 60% and nearly the same percentage of female students as male who scored above 85%. These results are reported in more detail in Lorenzo, Crouch, and Mazur.[40]

To see the effect of pedagogy on the gender distribution of final grades, we examined the final grade distribution of both genders in the calculus-based course at Harvard. While final grades are a somewhat unreliable measure of student learning because homework and class participation affect final grades, and because the format of exams and the method for calculating final grades changed from year to year over the study, grades still constitute the most important measure of success for the student. Figure 8 shows the histograms of the average percentage of each gender receiving grades of A, B, C, D, and E, in the (b) traditional, (c) hybrid, and (d) fully interactive years of the course. In all three graphs, the percentage of males receiving the highest grade of an A is consistently higher than the percentage of females; however, this gap reduces as interactivity increases. The distribution of female grades most closely matches that of the males in the fully interactive courses.

4. Using Peer Instruction at other institutions

Peer Instruction has been extensively used at Harvard University, and its success at increasing student understanding in those courses has been documented extensively. Informal conversation with faculty at many other institutions suggests that Peer Instruction has been very successful at a wide range of schools, from community colleges to large state universities to elite private colleges. We have recently undertaken a study of the effectiveness of PI at other institutions.

Through personal communication and the Project Galileo[41] database, we know of thousands of instructors who use PI in various colleges, universities, and secondary

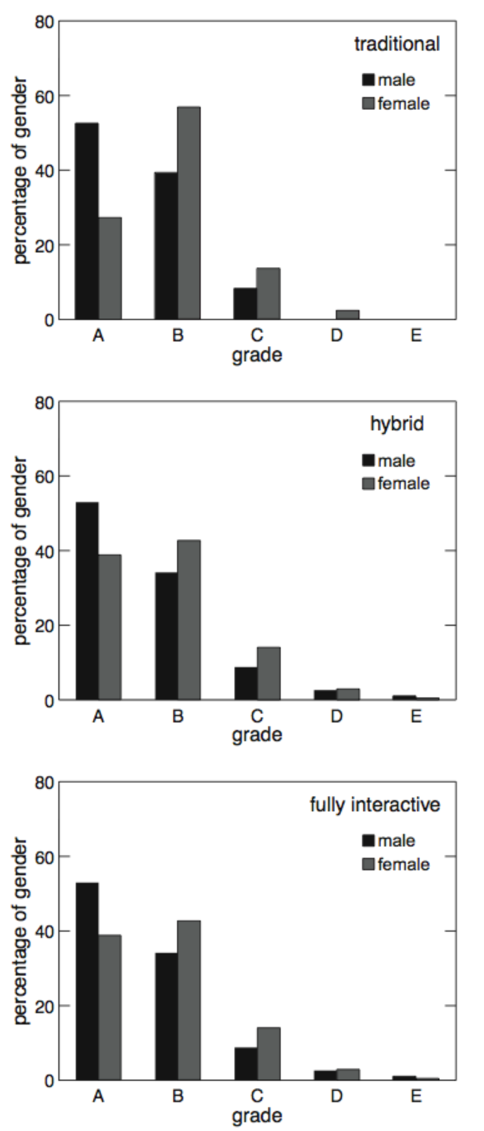


Fig. 8: Grade distribution by gender from introductory calculus-based physics course at Harvard University for the (*a*) traditionally taught course in 1990; (*b*) hybrid course using some interactive engagement methods in 1991, 1993, and 1995; (*c*) fully interactive course in 1996 and 1997.

schools around the world. In a study to examine PI at other institutions, we surveyed instructors on how they implement PI, and how effective it has proven to be in the variety of settings in which it is used.[7, 8, 9] Instructors using PI were asked to complete a web-based survey including questions on a variety of topics:

- Demographic information
- How they learned about Peer Instruction
- Course information
- Use of ConcepTests
- Grading and use of reading assignments
- Data on student understanding
- Student and instructor evaluation of PI
- Responses of colleagues

The survey was made available online at http://galileo.harvard.edu/Pisurvey.html (where it can still be found), and over 2700 instructors were invited by email to complete the survey. Over 700 instructors responded and completed the survey between June and December 1999. The language of the survey was purposely broad in order to include instructors who had used a strategy similar to PI; therefore data was also collected from respondents using other similar collaborative learning strategies. By looking at each respondent's description of his or her teaching methods, 384 of the respondents were classified as using Peer Instruction.[4, 6]

The sample was limited to those PI users that responded to our request to complete the survey, and therefore contains a selection bias, which may affect some of the responses. However, the respondents represent a broad array of institution types across the U.S. and around the world. With 23 countries represented, most responses were from the United States, Canada, and Australia. About two-thirds of survey respondents teach at universities (Figure 9), though almost all of their PI classes are for undergraduates, and typically introductory courses. The vast majority of respondents (82%) use PI to teach physics, although chemistry, life sciences, engineering, mathematics and astronomy courses are also represented. This finding is not surprising as the initial list of instructors contacted was biased towards those teaching physics and the pedagogy was developed originally to teach introductory undergraduate physics. There are also more materials available for physics (notably *Peer Instruc-*

tion: A User's Manual[4] and the Project Galileo[41] (with a database of ConcepTests primarily in introductory physics.)

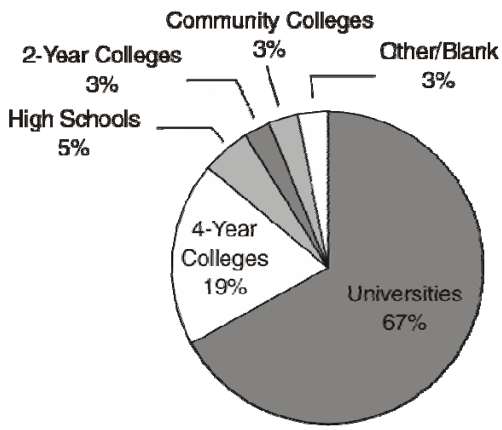


Fig. 9: Demographic breakdown of survey respondents using PI based upon institution (N = 384). Used by permission from ref. 8.

4.1. Effectiveness of Peer Instruction

The survey[9] probed three different measures of the success of PI in a course: data on student learning, student attitudes to the new pedagogy, and instructors' perception of the success of PI use.

Assessment of student learning

More than 100 PI users responding to the survey reported collecting quantitative data on the effectiveness of PI, of whom 81% administered the Force Concept Inventory[29] to assess their students' conceptual understanding of Newtonian mechanics. Instructors at 11 colleges and universities provided us with matched sets of pre- and post-FCI data, with data from 30 courses, to assess the gain for individual students. These PI courses have a class average normalized gain of 0.39 +- 0.09 (Figure 10). In his survey of FCI data, Hake[1] defines a "medium-g" range from g = 0.3 to 0.7 and finds that 85% of the interactive engagement courses included in his survey – and none of the traditionally taught courses – show gains in this range. We find that 27 out of 30 (90%) PI courses in our survey fall in the medium-g range with only three below g =

0.3 (Figure 10) This finding does not include data from Harvard courses, and shows the effects of PI when implemented by instructors other than the developers.

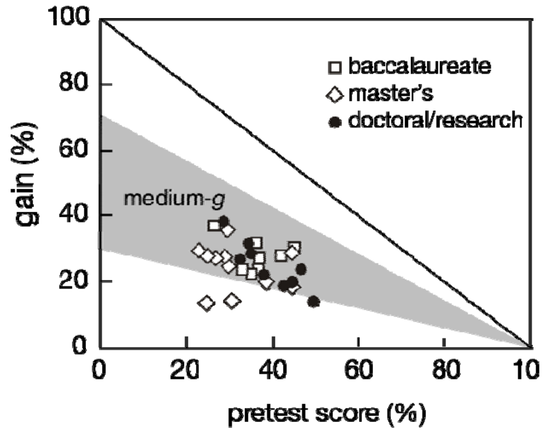


Fig. 10: Class-averaged normalized gain of introductory physics students in college courses taught using Peer Instruction. Symbols denote institution type according to the 2000 Carnegie Classification of Institutes of Higher Education. Used by permission from ref. 8.

We also asked all instructors to qualitatively report how the use of PI has affected their students' conceptual understanding of the material. Almost 60% of instructors who answered this question unambiguously perceive that student understanding has improved with PI, with an additional 20% stating that student understanding improved somewhat. Only 19% perceived no difference in student understanding from teaching traditionally, while 2% reported that they felt their students' understanding decreases with PI. Of those who perceived that their students' understanding suffered with PI, none collected any data.

Instructor satisfaction

The survey probed the instructor's opinion of PI and whether this experience proved to be valuable and/or enjoyable for them. The vast majority (93%) of respondents reported a positive experience. Additionally, to determine whether instructors consider the use of PI in their classes to be successful, instructors were asked if they were likely to use PI again in the future (Figure 11). Of the 384 identified PI users responding to the survey, 303 (79%) definitely planned to use PI again and 29 (8%) probably would. Only seven respondents (2%) expressed no plans to use PI again. Thus, the vast majority of instructors completing the survey consider their experiences with PI to be valuable and successful. These responses may not accurately represent the relative incidence of positive and negative experiences, as there may be a selection bias in who chose to complete the survey; however, the responses do indicate that PI has been successfully implemented in a large number of classrooms.

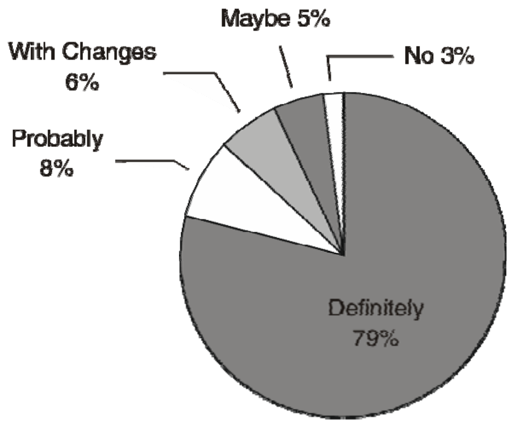


Fig. 11: Indication of instructors' plans to use PI in a future class (N = 384). Used by permission from ref. 9.

Student satisfaction

Instructors also report that their students are generally satisfied with PI, with 70% of instructors reporting more positive student evaluations and an additional 3% finding no difference from a traditional course. One university physics professor reported that while he observed similar performances on exams, student satisfaction was much higher: "the students almost universally enjoyed the class more and felt they understood more of what was being taught and, importantly, WHY it was being taught, *i.e.*, they saw the relevance more. The enjoyment level has been significantly higher (for them and for me!)"

Seventeen percent of instructors respond that students reported both positive and negative experiences, while approximately 5% of instructors report a dominant negative response from their students. Several of this latter group of instructors mention that this negative response occurred in spite of positive effects on student learning. According to one university physics professor, students responded that PI "was very useful…[but] they did not necessarily enjoy it." A 4-year college physics professor reports, "There was a lot of resistance to the idea [of PI]. They just wanted me to tell them the answer. Even at the end of the semester, after a test where groups that were working well together leaped way up on test performance, they still resisted trying to think for themselves." Students complained about having to think and actively participate in class, which is an activity that they are unaccustomed to doing. While it may be easier for some students to passively sit during lecture without being required to process the material, it is better for the students' learning to foster independent thinking in class.

Student retention

Survey respondents consistently report that the number of students who drop their PI course was dramatically lower than classes taught using a more traditional pedagogy. This is consistent with data from other studies, such as that of Williamson and Rowe[42] who report a 33.3% withdrawal rate from a traditional class, but only 17.3% from a treatment course in which lecture was replaced with cooperative group problem-solving.

4.2. How and where Peer Instruction is used

Course characteristics: who is teaching what with Peer Instruction?

Many questions on the survey were concerned with the nature of the classes taught with PI. In colleges and universities, the majority (90%) of courses using PI were introductory undergraduate courses. Class sizes varied significantly, ranging from 6 to 1200 students, with a mean of 122 students (standard deviation: 187). PI is being used successfully in classes of almost any size, from an intimate setting to the most immense lecture. The data collected in the survey reflects the education of over 45,000 students. This significantly underestimates the number of students taught using PI, however, as not only were we unable to survey all instructors using PI, but those that did complete the survey were asked to respond based on only their most representative or most recent *single* class. Many of the respondents had taught more than one separate course using PI and many had taught the same course with PI multiple times. It seems quite likely, therefore, that the actual number of students taught using PI is in the hundreds of thousands.

The majority of instructors were not new to their courses, and most had adapted their course to PI after teaching it in another way previously. In fact, 72% of instructors had taught the same course previously using the lecture format and 13% had used some other format for the same course. It is important to note that even those who presumably had a set of prepared lecture notes found the switch to PI to be worthwhile.

Implementation details: how instructors use Peer Instruction

Peer Instruction is an inherently flexible and adaptable method. In order to determine how extensively instructors adapt PI, the survey asked many questions about details of implementation. In analyzing the survey responses, we were particularly interested in learning whether other instructors find that non-competitive grading in the course and pre-class reading are as important as we have found them to be at Harvard. As discussed below, it appears that PI is adapted extensively to match what instructors believe to be the needs of their students. We are presently beginning a multiyear study to assess whether these implementation variations affect student learning (either positively or negatively).

Grading and competitiveness

As mentioned in Section 2, classes at Harvard incorporate an absolute grading scale to establish a non-competitive atmosphere within the classroom.[4] This atmosphere is necessary so that students do not fear they will jeopardize their own grades by working with their classmates. Over 59% of our survey respondents reported grading on an absolute scale. However, nearly a third of respondents do grade on a curve (32%), and some use a different grading strategy (16%). Other respondents report using more than one grading scheme.

Preclass reading

As mentioned in Section 2, we found pre-class reading and assignments following up on the reading to be very helpful in our courses at Harvard.[4, 6] Therefore, we were interested to see if and how pre-class reading assignments were implemented in others' courses. Over 71% of respondents using PI report that they also expect pre-class reading from their students. However, nearly 25% of those requiring reading do not have any means for assessing whether that reading had actually been completed (Table 3), so it is uncertain how many students are actually completing the reading before class. Combining this with the approximately 30% of all instructors who do not require reading at all, our results show that nearly 45% of all PI users responding to our survey either do not require pre-class reading or do nothing to assess their students' compliance with any reading assignments. We thus found that many faculty do not appear to consider pre-class reading assignments essential to PI.

The reading assessments that were most common were in the form of a multiple choice or free-response reading quiz, typically administered before class. In contrast to our classes at Harvard, the majority of instructors that incorporate reading assignments (71%) assign students a grade based on the content and correctness of their answers. Approximately one-third of instructors follow a similar grading strategy to the one we have used at Harvard, by assigning a grade based on the effort shown or giving credit for simply completing the assignment.

Table 3: Type of reading assessment used by those instructors requiring pre-class reading (nearly 30 % of respondents did not require such reading). Indicates the percentage of respondents who use the listed method (*N* = 274). Respondents could choose more than one option. Used by permission from ref. 9.

Type	% using
Multiple-choice reading quiz	52%
Free-response reading quiz	17%
Reading summary	8%
Other assessment	17%
No reading assessment	25%

Polling for student responses

As discussed in Section 2, there are several methods that can be used in polling for student responses to ConcepTests. Among respondents to our survey, 65% ask for a show of hands, and about a third use flashcards. Only 8% of survey respondents report using an electronic polling method, and these classes tend to be large (average enrollment of 208, versus 122 for all classes represented in the survey). Many instructors reported using multiple polling methods. It seems likely that the percentage of faculty using such electronic polling devices has increased since the survey was conducted in 1999, as such systems are becoming far more common and even are packaged with some textbooks.

4.3. How faculty respond to the challenges faced in teaching with Peer Instruction

Many instructors who perceive Peer Instruction as a successful teaching strategy also indicate that they had to overcome a number of challenges. Thirteen percent of instructors cite the time and energy required to develop ConcepTests as an impediment to using PI. Developing good ConcepTests certainly takes a great deal of effort; to minimize duplication of this effort, and to make PI easier to implement, we and other developers of ConcepTests have made online databases for introductory physics, chemistry, and astronomy freely available, as described in Section 5.

Ten percent of respondents report that their colleagues are skeptical of the benefit of student discussions that take away from lecture time. A third of these instructors report addressing this skepticism by collecting data on student learning gains. To support instructors in doing so, the Interactive Learning Toolkit (described in Section 5) makes it possible for instructors to administer the Force Concept Inventory[29] and Mechanics Baseline Test[30] through a secure online interface.[41] Another approach found effective by instructors is to compare achievement of students taught with and without PI on identical exams.[4] Others suggest inviting skeptical colleagues to sit in on a class, sharing positive student feedback with them, or even giving tests of conceptual understanding to other faculty.

About 9% of respondents report that the quantity of material that they must cover in a semester makes it difficult to devote class time to ConcepTests. One-tenth of these instructors reduce the amount of material covered by the course, but most do not have the freedom to do so. While we have had success at Harvard and Swarthmore requiring students to learn some of the material from reading on their own, some instructors may encounter significant resistance to this from students, or may feel that some of their students are not sufficiently independent or sophisticated to learn this way.

Another challenge is students' resistance to the method (7% of respondents). As mentioned previously, students are typically unaccustomed to actively participating in science classes, so it is not surprising that some feel uncomfortable participating in discussions, or initially consider the discussions a waste of time. Thus, respondents report, and confirm our arguments, that it is essential to thoroughly explain the use of PI to their students. Techniques for this motivation were highlighted in Section 2 of this article, as well as in *Peer Instruction: A Manual*.[4] Persistence in the face of initial student resistance is also important; 15 (4%) users report that, while their students were initially skeptical of PI, the students warmed up to it as they found the method helped them learn the material. Regularly presenting class-averaged data on student performance also shows students that the method is helping them may also motivate students.

A related challenge is the difficulty in fully engaging students in class discussions (7% of respondents). In the words of one instructor, "some students were too cool, too alienated, or perhaps too lost to participate." Nearly half of those citing this challenge say it is important for the instructor to circulate through the classroom during the group discussion of the ConcepTest, helping to guide and encourage students in discussion. Other students may be motivated by receiving credit for participation and by the presence of ConcepTest-like conceptual questions on exams[4] as described in Section 2.

5. Resources available for teaching with Peer Instruction

Many resources available for implementing PI in introductory physics courses (as well as in mathematics, chemistry, and astronomy courses). As the focus of education shifts to target students' conceptual understanding, the resources for implementing PI are growing to include disciplines such as biology, economics, and psychology, as well as advanced undergraduate science courses.

We begin this section of the article by providing all the components of a sample class — reading questions with representative student answers, lecture outline, ConcepTests, associated assignments, and possible exam questions to ask on this topic. Next, we provide information on where to find ConcepTests on other topics. The following section describes the Interactive Learning Toolkit, a web-based course management system designed specifically to support teaching with PI, and finally, we list other useful resources for those teaching with PI.

5.1. An example: Materials for a class introducing electric fields

As an example of how teaching with PI works, we describe here the process of teaching students about electric fields, including a 90-minute class on the topic. Before coming to class, students are required to read the appropriate section of their textbook.[31] The reading covers the following topics:

1) The field model
2) Electric field diagrams
3) Superposition of electric fields
4) Electric fields and forces

By midnight the evening before class, students are required to complete the reading assignment on the Interactive Learning Toolkit.[41] The instructor then reviews student responses to three questions, two that test students understanding of the concepts and another that asks them about what they found confusing, to develop the lecture outline and ConcepTests needed for class. The questions and sample student responses to this assignment are listed below.

Question 1: Two charged particles are held fixed a certain distance apart. If the electric field vector is zero at some point on the line between the two charges, what can we conclude about the signs of the charge on the particles? About the magnitudes of their charges? Briefly explain your reasoning.

Answer (provided to the students on the web site the day after the assignment is due): If the electric field is zero anywhere between the two charges, it implies that the sign of the two charges is the same. The relative magnitudes of the charges cannot be determined from the information given. The point of zero field will be closer to the more weakly charged particle.

Almost all of the students correctly identified the signs of the charges; however, many believed that the magnitudes of the charges must be equal as well.

> *Sample student answers:*
>
> The signs of the charge must be the same. A third charge at the point where the electric field is zero must be either repelled by both of the charges or attracted by both of them, which can only occur if these charges are of the same sign. We can't conclude anything about the magnitude of the charges, because we don't know how close the "zero point" is to each charge respectively.
>
> We can conclude that the signs of the charge on both particles is the same and both are of equal magnitude, allowing for the electric field vector to be zero as a result of the vector sum of the forces exerted by both particles to cancel each other out.
>
> If the electric field is zero anywhere between the two charges, it implies that the sign of the two charges is the same. The relative magnitudes of the charges cannot be determined from the information given. The point of zero field will be closer to the more weakly charged particle.

Question 2: A charged particle is released from rest in a region of uniform electric and gravitational fields. The electric field is directed perpendicular to the gravitational field, and the acceleration of the particle due to the electric field is comparable in magnitude to that due to the gravitational field. Describe the trajectory of the particle in the combined electric and gravitational fields.

Answer: The particle is subject to a downward gravitational force and a horizontal electrostatic force. The vector sum of these forces is directed diagonally downward (at about 45 degrees if the two forces are of roughly equal magnitude) and so it is accelerated in that direction. Because the particle starts at rest, its trajectory is a straight line in the direction of the vector sum of the forces.

Most students were able to add the forces correctly; however, some students had some misconceptions about vector sums. Several students indicated that the particle followed a curved or parabolic trajectory.

> *Sample student answers:*

Because the accelerations due to each field are equal in magnitude and at a 90 degree angle to one another, we can assume the trajectory is a straight line at 45 degrees (combining the vectors).

If the accelerations due to the electric field and gravitational field are similar in magnitude, the particle should follow a diagonal trajectory, 45 degrees between the electric and gravitational fields.

The particles will have a parabolic trajectory in which the particles will "fall" vertically from the gravitational field and move in a horizontal path from the electric field. The exact direction depends on the charge of the particle. If it is positive then the force and accelaration [sic] of the electric field and if the charge is positive the accelaration will point opposite to the electric field.

The trajectory of the particle would be a sharp curve downward. (I really have no idea...)

Question 3: Please tell us briefly what **single** point of the reading you found most difficult or confusing. If you did not find any part of it difficult or confusing, please tell us what parts you found most interesting.

Sample student answers:

I'm confused about the dipole moment vector. If a particle with a dipole moment enters a uniform electric field, you say that the dipole has zero acceleration. Does this mean that the particle doesn't move in the field at all?

I am getting the force and field vectors confused. I don’t really understand the difference, and I can’t differentiate when I should be looking at the effect of one or another on a particle. Do the force vectors determine the field vector? Are they completely independent of each other?

I didn't understand the trajectory of particles, I have a hard time visualizing what is happening.

> I did not understand the idea of torque in the electric field and how to determine the effects of torque or in what direction it moves the particles. Does torque move the particles or the electric field, or both?

The students' responses indicate that the main conceptual difficulties are concerned with:

1) Difference between electric field and force
2) Dipoles in uniform and nonuniform fields
3) Trajectories
4) Torques on dipoles

During lecture, the instructor begins by reviewing the key concepts that have already been introduced in the reading. Coulomb's law is written on the board along with a sketch showing charges q_1 and q_2 separated by a distance r. Using the equation, the instructor goes on to describe the difference between an electric field at point 2 and the force on a charge at point 2. After presenting fields and forces, the ConcepTest shown in Figure 12 is used to check students' understanding of how to use forces to measure the electric field. The distribution of student answers indicates that just over half of the students answered correctly. Therefore, a discussion between students is necessary to help solidify the concept for those who answered correctly and gives the opportunity for those who answered incorrectly to clear up their misconceptions about electric fields. After a few minutes of discussion, during which the teaching assistants and instructor circulate and listen, students submit their revised answers and the percentage of correct answers jumps to 68%. A few more minutes are spent explaining the correct answer to the ConcepTest.

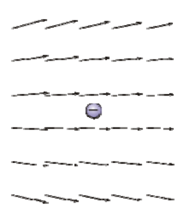
To follow up on the concept, the ConcepTest shown in Figure 13 is presented. An overwhelming majority (90%) of students were able to answer this question correctly; therefore there is no need for student discussion. The instructor reviews the answer briefly with the class and summarizes what they have learned about fields.

The lecture moves to the next point: dipoles. After reviewing the definition, the instructor asks students to respond to the ConcepTest in Figure 14. Before discussion, You are given a 'test' particle to map out the electric field of a charged object. To correctly determine the object's electric field, you need to know:

1. the magnitude and sign of the charge on the test particle.
2. the magnitude and sign of the charge on the object.
3. all of the above.
4. only the sign of the charge on the test particle.
5. None of the above.

Fig. 12: First ConcepTest on electric fields. Student responses before discussion were (1) 56%, (2) 24%, (3) 12%, (4) 5%, (5) 3% and after discussion were (1) 68%, (2) 22%, (3) 5%, (4) 1%, (5) 3%.

A negatively charged object is placed in an electric field as shown below.



The direction of the electrostatic force on the object is

1. to the right
2. to the left
3. neither to the left nor to the right
4. depends on whether the field is created by a positively or negatively charged object
5. There is no force on the object at the location shown in the figure

Fig. 13: Second ConcepTest on electric fields. Student responses before discussion were (1) 6%, (2) 89%, (3) 0%, (4) 6%, (5) 0%.

about 70% of students were able to answer the question correctly. During discussion, the teaching assistants encourage students to engage with someone that disagrees with their answer. After discussion, almost all (97%) of the students gave the correct answer. The answer is briefly explained with a few diagrams showing the vector sums.

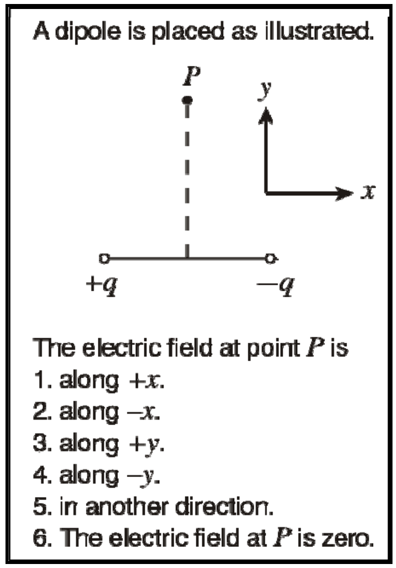


Fig. 14: First ConcepTest on dipoles. Student responses before discussion were (1) 70%, (2) 4%, (3) 8%, (4) 6%, (5) 4%, (6) 9% and after discussion were (1) 97%, (2) 1%, (3) 0%, (4) 0%, (5) 1% (6) 0%.

The final ConcepTest is also concerned with dipoles and is shown in Figure 15. Very few students (14%) were able to answer this question correctly. Before giving the students a chance to talk to their neighbor, the instructor made a sketch for students to give them a visual image of the question before thinking about it further and also briefly reviewed the ideas of uniform and nonuniform fields. With a little help and discussion with their neighbors, about 75% submitted correct revised answers. The instructor drew a diagram on the board depicting the forces on a permanent dipole in both uniform and nonuniform fields.

With just a few minutes remaining, the instructor reminded students that they can access these ConcepTests and one other (Figure 16) relating to electric fields on the website. If they have confusion with these, an online forum is provided on the Interactive Learning Toolkit[41] for them to pose questions.

Does a neutral object experience a force in an electric field?

1. No
2. Only in a uniform field
3. Only in a nonuniform field
4. Yes, because the field polarizes the object

Fig. 15: Second ConcepTest on dipoles. Student responses before discussion were (1) 26%, (2) 1%, (3) 14%, (4) 57% and after discussion were (1) 9%, (2) 1%, (3) 75%, (4) 15%.

Does a neutral object have an electric field? (Consider points outside the object.)

1. Yes.
2. No.
3. It depends.

Fig. 16: ConcepTest (not used in class) on uniform and nonuniform electric fields.

After lecture, students have further opportunities to practice with their conceptual understanding of electric fields. In discussion sections led by a teaching fellow, groups of three or four students work through the “Charge” tutorial in Tutorials in Introductory Physics.[3] In addition, they are assigned a problem set (Appendix 3) that consists of more quantitative problems.

5.2. Other sources of ConcepTests

The hallmark of Peer Instruction is the ConcepTest, and the key to implementing PI is in developing good conceptual questions to probe students understanding. Already there are a number of resources providing class-tested, ready-to-use conceptual questions. *Peer Instruction: A User’s Manual*[4] includes 243 ConcepTests developed for our introductory calculus-based physics for non-majors, covering mechanics, electricity, magnetism, fluid statics and dynamics, oscillations and waves, geometrical and physical optics, and modern physics. A collection of ConcepTests in chemistry[43] has also been published and includes conceptual questions in general, inorganic, and organic chemistry, drawn from an online database maintained at the University of Wisconsin. For introductory astronomy, Paul Green has compiled ConcepTests covering a large range of topics.[44] In calculus, an instructor's supplement with ConcepTests was developed to accompany *Calculus* (3rd edition) by Hughes Hallett *et al*.[45]

A searchable database of ConcepTests on the Interactive Learning Toolkit website, formerly the Project Galileo Web site (http://www.deas.harvard.edu/ilt; 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 non-introductory 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.

5.3. Interactive Learning Toolkit (online course management system specifically designed to facilitate teaching with Just-in-Time Teaching and Peer Instruction)

In addition to providing a database of ConcepTests, the ILT also can be used for online course management, with tools to help organize and design courses to link lectures, reading assignments, question and answer forums, and in-class participation to ConcepTests. The ILT was created to encourage increased communication between the instructor and the students, as well as between the students and their understanding of the concepts in the course. A brief description of the modules and their functions are below:

- *Lectures* can be used to design each class meeting and are linked by dates and times. With a database of ready-to-use ConcepTests, the instructor can choose which conceptual questions may best probe students' understanding of the reading, lecture, assignments, and content. Additionally, the ILT provides an easy way to create additional ConcepTests in pdf format, which can also be shared and added to the database. The instructor can easily generate a slide set of all ConcepTests that may be useful for a given lecture topic, and post these for students to access after class.

- If the instructor uses the *Peer Response System* to poll students for answers to in-class ConcepTests, the *Lecture* module of the ILT contains a feature to link responses with the students. This is useful for analyzing student answers with other areas of the course, such as performance on pre-class reading, assignments, and exams. Additionally, we have integrated the technology of the ILT with BQ[46] to allow students to use wireless-enabled devices, such as cell phones, laptops, or PDAs, to respond to in-class ConcepTests. With many students already using these devices in class, this feature alleviates the need for students to purchase an additional device and reduces the technical infrastructure needed in the classroom.

- The *Reading* module provides features to help create and post pre-class reading assignments. These assignments are anchored to the dates and times of the lecture. Students are able to see the assignment and its due date. Instructors and teaching assistants are able to quickly review all student responses to a given question, revealing common weaknesses in the class's understanding. Therefore, ConcepTests can be chosen in advance to specifically target and probe these areas of the content. The ILT also permits instructors to respond to questions or misconceptions expressed in student responses via a labor-saving web interface, increasing students' sense of individual connection to the instructor.

- *Assignments* can be used to post homework or problem sets that need not be linked to the lectures or reading assignments. The ILT has built-in features for due dates and reminders for students, and enables the instructor to post solutions at specific times. As with most online course management systems, the ILT provides an online gradebook for instructors to use and students to access.

- *Standardized Tests*, including those mentioned in this article, are available on the ILT and can be provided to students as online assignments. These tests are designed to assess students' conceptual understanding, quantitative problem-solving skills, or attitudes about undergraduate science courses, and can be taken pre- and post-course to provide information on the effectiveness of the instruction in these specific areas. The database of these tests is growing, and currently include the Force Concept Inventory,[29] Mechanics Baseline Test,[30] Astronomy Diagnostic Test,[47] Conceptual Survey on Electricity and Magnetism,[48] Lawson's Test of Scientific Reasoning,[49] and the Maryland Physics Expectations Survey.[24] Other standardized tests can be easily added to the database.

- *Forum* provides a place for online discussions between the instructor(s), teaching assistants, and students. Students can post administrative and content-related questions that could be of benefit to the entire class. The forum also enables students to form an online learning community to discuss course concepts and further collaboration. The instructor and teaching assistants can initiate and monitor these discussions, providing another convenient way to observe students' understanding of the key concepts of the course.

- *News* and *Handouts* are additional features, common to web-based course management systems, which allow the instructor to post announcements and updates, as well as files and other course material for students to access.

The ILT software is freely available to any interested instructor, requiring only that the instructor register at http://www.deas.harvard.edu/ilt. (In order to preserve the security of standardized tests, such as the Force Concept Inventory, an instructor must also send email to galileo@deas.harvard.edu in order to gain access to these tests.) The ILT site includes a user's manual and a quick-start manual for using the standardized test feature of the site.

5.4. Other publicly available resources for teaching with Peer Instruction

The Peer Instruction User's Manual,[4] in addition to providing ConcepTests as described above, provides a set of conceptual exam questions covering the typical topics of an introductory two-semester physics course.

For new faculty members, Eric Mazur regularly attends the AAPT New Faculty Workshop to demonstrate Peer Instruction as well as answer questions and provide advice for users that have attempted PI in their classes already. Additional support can be obtained by communicating directly with the education members of the Mazur group, found on the People page at http://mazur-www.harvard.edu.

As mentioned previously, we have had the greatest success with Peer Instruction when coupled with elements of Just-in-Time Teaching, the Tutorials in Introductory Physics, and Cooperative Group Problem Solving. All of those methods are described more extensively in other articles in this volume.

6. Conclusions

Peer Instruction has been used successfully at hundreds of institutions around the world and has produced substantial gains in student understanding at Harvard University, where it has been most extensively evaluated. Learning gains are greatest when PI is complemented by other strategies that increase student engagement, so that every element of the course serves to involve students actively. At other institutions, PI has been shown to produce Force Concept Inventory gains commensurate with Hake's findings on learning gains obtained with interactive engagement methods. If significant effort is invested in motivating students, student reactions to PI are generally positive, though there are always some students resistant to being taught in a non-traditional manner. Finally, extensive resources are available for teaching with Peer Instruction in introductory physics courses, as well as for other courses, and

Peer Instruction is inherently adaptable to an instructor's particular classes and students.

Acknowledgements

The authors would like to thank Prof. Michael J. Aziz, Prof. William Paul, Prof. Chris Schaffer, Dr. Paul Callan, and other members of the Physics 1 and 11 teaching staff at Harvard University for their contributions to the development of Peer Instruction and ConcepTests; Prof. David Sokoloff and Prof. Michael Wittman for helpful discussion of the demonstration study; Prof. Stamatis Vokos for helpful discussion of the comparison of ConcepTests and exam questions. Development of Peer Instruction and its accompanying resources was funded by a series of grants from the National Science Foundation and the Division of Applied Sciences at Harvard University.

References

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

[2] D.W. Johnson, R.T. Johnson, and K.A. Smith, *Active Learning: Cooperation in the College Classroom* (Interaction Book Company, 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, 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, 1998); P. Laws, *Workshop Physics* (John Wiley & Sons, 2001); D. K. Sokoloff and R. K. Thornton, *Interactive Lecture Demonstrations* (John Wiley & Sons, 2004) (described in D. K. Sokoloff and R. K. Thornton, *Phys. Teach.* **35**, 340 (1997)); A. Van Heuvelen and E. Etkina, *The Physics Active Learning Guide,* (Pearson Addison Wesley, 2006.) and numerous forms of Socratic dialogue, as in R.R. Hake, "Socratic Pedagogy in the

Introductory Physics Lab," *Phys. Teach.* **30**, 546 (1992), or group problem solving, as in P. Heller, R. Keith, and S. Anderson, "Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving," *Am. J. Phys.* **60**, 627 (1992), and P. Heller and M. Hollabaugh, "Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups," *Am. J. Phys.* **60**, 637 (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, 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," *Optics & Photonics News* **9**(9), 37 (September 1998).

[6] Catherine H. Crouch and Eric Mazur, "Peer Instruction: Ten Years of Experience and Results," *Am. J. Phys.* **69**, 970 (2001).

[7] A. P. Fagen, C. H. Crouch, T. Yang, and E. Mazur, "Factors That Make Peer Instruction Work: A 700-User Survey," talk given at the 2000 AAPT Winter Meeting, Kissimmee, FL, January 2000.

[8] A. P. Fagen, C. H. Crouch, and E. Mazur, "Peer Instruction: Results From A Range of Classrooms," *Phys. Teach.* **40**, 206 (2002).

[9] A. P. Fagen, "Assessing and Enhancing the Introductory Science Course in Physics and Biology: Peer Instruction, Classroom Demonstrations, and Genetics Vocabulary," Ph.D. thesis, Harvard University (2003), and references therein.

[10] C. H. Crouch, A. P. Fagen, P. Callan and E. Mazur, "Classroom Demonstrations: Learning Tools or Entertainment?" *Am. J. Phys.* **72,** 835 (2004).

[11] D. E. Meltzer and K. Manivannan, "Promoting interactivity in physics lecture courses," *Phys. Teach.* **34**, 72 (1996).

[12] More information about the Peer Response System can be found at http://www.educue.com.

[13] G. Novak, E. Patterson, A. Gavrin, and W. Christian, *Just-in-Time Teaching: Blending Active Learning and Web Technology* (Prentice Hall, 1999), and http://webphysics.iupui.edu/jitt/jitt.html.

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

[15] S. 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.

[16] 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]

[17] Questions on the introductory calculus-based physics final exam were categorized as conceptual or quantitative, based on the skills that the question assessed. The average score on the conceptual questions was 3.24 (sd=1.25) and the average score on the quantitative questions was 3.52 (sd=1.39). A *t*-test (two-tailed) was performed to determine the likelihood that the difference in scores is due to real differences in student performance on the two types of question. With a p-value of < 0.0001, the difference in the average scores was statistically significant.

[18] P. M. Sadler, “Psychometric Models of Student Conceptions in Science: Reconciling Qualitative Studies and Distractor-Driven Assessment Instruments,” *J. Res. Sci. Teach*. **35**, 265 (1998), and “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 Prof. Sadler (private communication).

[19] R. M. Felder and R. Brent, “Navigating the Bumpy Road to Student-Centered Instruction,” *College Teaching* **44**, 43 (1996).

[20] 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 (1994); and Felder and Brent [ref. 19].

[21] D.W. Bullock, V. P. La Bella, T. Clingan, Z. Ding, G. Stewart, and P. M. Thibado, “Enhancing the Student-Instructor Interaction Frequency,” *Physics Teacher* **41**(5), 272 (2002).

[22] 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, $N = 789$).

[23] 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 ($N = 229$).

[24] E. F. Redish, J. M. Saul, and R. N. Steinberg, "Student Expectations in Introductory Physics," *Am. J. Phys.* **66**, 212 (1998).

[25] L. R. Jones, J. F. Watts and A. G. Miller, "Case Study of Peer Instruction in Introductory Physics Classes at the College of Charleston," submitted to the proceedings of *Charleston Connections: Innovations in Higher Education*, 2000.

[26] N. Ambady and R. Rosenthal, "Half a Minute: Predicting Teacher Evaluations From Thin Slices of Nonverbal Behavior and Physical Attractiveness," *J. Personality and Social Psych.* **64**, 431 (1993).

[27] 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.

[28] 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.

[29] 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 (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 Prof. 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.

[30] D. Hestenes and M. Wells, "A Mechanics Baseline Test," *Phys. Teach.* **30** (3), 159 (1992). This test is available from the same sources as the FCI [ref. 29].

[31] 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.

[32] 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.

[33] 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 Prof. Laura McCullough of the University of Wisconsin–Stout for telling us about this survey.

[34] The *p*-value was 0.26; a *p*-value of 0.05 or less is generally agreed to indicate a statistically significant difference.

[35] 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.

[36] The questions we identified as significantly quantitative are numbers 9, 11, 12, 17, 18, 23, 24, and 25 (8 in all).

[37] 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 (1994).

[38] Catherine H. Crouch, J. Paul Callan, Nan Shen, and Eric Mazur, “ConcepTests: What do students learn from them?” talk given at the 2000 AAPT Winter Meeting, Kissimmee, FL, January 2000.

[39] E. Seymour and N. Hewitt, *Talking about Leaving: Why Undergraduates Leave the Sciences* (Westview Press, 1997); B. L. Whitten, S. R. Foster, and M. L. Duncombe, "What works for women in undergraduate physics?" *Physics Today* **56** (9), 46 (Sept. 2003); and references therein.

[40] M. Lorenzo, C. H. Crouch, and E. Mazur, "Reducing the gender gap in the physics classroom," *Am. J. Phys.* **74**, 118 (2006).

[41] Project Galileo is an online server, providing resources for innovative teaching ideas and materials. Much of the site has since been incorporated into the Interactive Learning Toolkit, an online course management system, described in detail in Section 2 and accessible at http://www.deas.harvard.edu/ilt.

[42] V. M. Williamson and M. W. Rowe, "Group problem-solving versus lecture college-level quantitative analysis: the good, the bad, and the ugly," *J. Chem. Ed.* **79,** 592 (2002).

[43] C. R. Landis, A. B. Ellis, G. C. Lisensky, J. K. Lorenz, K. Meeker, and C. C. Wamser, *Chemistry ConcepTests: A Pathway to Interactive Classrooms* (Prentice Hall, 2001).

[44] P. Green, *Peer Instruction for Astronomy* (Prentice Hall, 2002).

[45] D. Hughes-Hallet, et al., *Calculus: ConcepTests.* (John Wiley and Sons, 2003).

[46] The BQ software has been integrated with the ILT to allow students to respond to multiple-choice questions via networked desktop, laptop computers, web enabled cell phones, wireless PDAs, and/or wireless IR keypads. http://www.erskine.edu/bq/.

[47] B. Hufnagel, T. Slater, et al., "Pre-course Results from the Astronomy Diagnostic Test," *Publications of the Astronomical Society of Australia*, **17**, 152 (2000).

[48] D. P. Maloney, T. L. O'Kuma, C. J. Hieggelke, and A. Van Heuvelen, "Surveying students' conceptual knowledge of electricity and magnetism," *Am. J. Phys.* **69,** S12 (2001).

[49] A. E. Lawson, "The development and validation of a classroom test of formal reasoning," *J. Res. Sci. Teach.* **15** (1), 11 (1978).