Confessions of a converted lecturer*

Eric Mazur
Department of Physics and School of Engineering and Applied Sciences, Harvard University, 9 Oxford Street, Cambridge, Massachusetts 02138, USA

The introductory physics course often is one of the biggest hurdles in the academic career of a student. For a sizable number of students, the course leaves a permanent sense of frustration. I have only to tell people I am a physicist to hear grumblings about high school or college physics. This general sense of frustration with introductory physics is widespread among non-physics majors required to take physics courses. Even physics majors are frequently dissatisfied with their introductory courses, and a large fraction of students initially interested in physics end up majoring in a different field. What have we done to make it that way, and can we do something about it? Or should we just ignore this phenomenon and concentrate on teaching the successful student who is going on to a career in science?

An eye opener

Frustration with introductory physics courses has been commented on since the days of Maxwell and has recently been widely publicized by Sheila Tobias, who asked a number of graduate students in the humanities and social sciences to audit introductory science courses and describe their impressions. [1] The result of this survey is a book that paints a bleak picture of introductory science education. One may be tempted to brush off

* This paper, which accompanies a lecture delivered in May 2007 in Oporto, Portugal, is adapted from *Peer Instruction: A User's Manual* (Prentice Hall, 1997) by Eric Mazur.

complaints by non-physics majors as coming from students who are *a priori* not interested in physics. Most of these students, however, are not complaining about other required courses outside their major field. In science education, in Tobias' words, the next generation of science workers is expected to rise like cream to the top, and the system is unapologetically competitive, selective, and intimidating, designed to winnow out all but the top tier.

The way physics is taught now is not much different from the way it was taught — to a much smaller and more specialized audience — a century ago, and yet the audience is vastly changed. Physics has become a building block for many other fields, and enrollment in physics courses has grown enormously, with the majority of students not majoring in physics. This shift in constituency has caused a significant change in student attitude toward the subject and made the teaching of introductory physics a considerable challenge. Although conventional methods of physics instruction have produced many successful scientists and engineers, far too many students are unmotivated by the conventional approach. What, then, is wrong with it?

I have been teaching an introductory physics course for engineering and science majors at Harvard University since 1984. Until 1990 I taught a conventional course consisting of lectures enlivened by classroom demonstrations. I was generally satisfied with my teaching — my students did well on what I considered difficult problems, and the evaluations I received from them were very positive. As far as I knew, there were not many problems in *my* class.

In 1990, however, I came across a series of articles by Halloun and Hestenes [2–5] that really opened my eyes. As is well known, students enter their first physics course possessing strong beliefs and intuitions about common physical phenomena. These notions are derived from personal experience and color students' interpretations of material presented in the introductory course. Halloun and Hestenes show that instruction does little to change these 'common-sense' beliefs.

For example, after a couple of months of physics instruction, all students can recite Newton's third law and most of them can apply it in numerical problems. A little probing, however, quickly shows that many students do not understand the law. Halloun and Hestenes provide many examples in which students are asked to compare the forces exerted by different objects on one another. When asked, for instance, to compare the forces in a collision between a heavy truck and a light car, many students firmly believe the heavy truck exerts a larger force. When reading this, my first reaction was 'Not *my* students…!' Intrigued, I decided to test my own students' conceptual understanding, as well as that of the physics majors at Harvard.

The first warning came when I gave the Halloun and Hestenes test to my class and a student asked, 'Professor Mazur, how should I answer these questions? According to what you taught us, or by the way I *think* about these things?' Despite this warning, the results of the test came as a shock: the students fared hardly better on the Halloun and Hestenes test than on their midterm examination. Yet, the Halloun and Hestenes test is *simple*, whereas the material covered by the examination (rotational dynamics, moments of inertia) is, or so I thought, of far greater difficulty.

Memorization versus understanding

To understand these seemingly contradictory observations, I decided to pair, on the examinations, simple qualitative questions with more difficult quantitative problems on the same physical concept. An example of a set of such questions on dc circuits is shown in Fig. 1. These questions were given as the first and last problem on a midterm examination in the spring of 1991 in a conventionally taught class (the other three problems on the examination, which were placed between these two, dealt with different subjects and are omitted here).

Note that question 1 is purely conceptual and requires only knowledge of the fundamentals of simple circuits. Question 5 probes the students' ability to deal with the same concepts, now presented in the conventional numerical format. It requires setting up and solving two equations using Kirchhoff's laws. Most physicists would consider question 1 easy and question 5 harder. As the result in Fig. 2 indicates, 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 serious misconception, many still managed to correctly solve the mathematical problem.

Figure 3 shows the lack of correlation between scores on the conceptual and conventional problems of Fig. 1. Although 52% of the scores lie on the broad diagonal band, indicating that these students achieved roughly equal scores (±3 points) on both questions, 39% of the students did substantially worse on the conceptual question. (Note that a number of students managed to score zero on the conceptual question and 10 on the conventional one!) Conversely, far fewer students (9%) did worse on the conventional question. This trend was confirmed on many similar pairs of problems during the remainder of the semester: students tend to perform significantly better on standard textbook problems than on conceptual ones covering the same subject.

This simple example exposes a number of problems one faces in science education. First, it is possible for students to do well on conventional problems by memorizing algorithms without understanding the underlying physics. Second, as a result of this, it is possible for a teacher, even an experienced one, to be completely misled into thinking that students have been taught effectively. Students are subject to the same misconception: they believe they master the material and then are severely frustrated when they discover that their plug-and-chug recipe doesn't work in a different problem.

Clearly, many students in my class were concentrating on learning 'recipes,' or 'problem-solving strategies' as they are called in textbooks, without considering the underlying concepts. Plug and chug! Many pieces of the puzzle suddenly fell into place:

- The continuing requests by students that I do more and more problems and less and less lecturing — isn't this what one would expect if students are tested and graded on their problem-solving skills?
- The inexplicable blunders I had seen from apparently bright students — problem-solving strategies work on some but surely not on all problems.
- Students' frustration with physics — how boring physics must be when it is reduced to a set of mechanical recipes that do not even work all the time!

One problem with conventional teaching lies in the presentation of the material. Frequently, it comes straight out of textbooks and/or lecture notes, giving students little incentive to attend class. That the traditional presentation is nearly always delivered as a monologue in front of a passive audience compounds the problem. Only exceptional lecturers are capable of holding students' attention for an entire lecture period. It is even more difficult to provide adequate opportunity for students to critically think through the arguments being developed. Consequently, all lectures do is reinforce students' feeling that the most important step in mastering the material is solving problems. The result is a rapidly escalating loop in which the students request more and more example problems (so they can learn better how to solve them), which in turn further reinforces their feeling that the key to success is problem solving.

Why lecture?

The first time I taught introductory physics, I spent much time preparing lecture notes, which I would then distribute to my students at the end of each lecture. The notes became popular because they were concise and provided a good overview of the much more detailed information in the textbook.

Halfway through the semester, a couple of students asked me to distribute the notes in advance so they would not have to copy down so much and could pay more attention to my lecture. I gladly obliged, and the next time I was teaching the same course, I decided to distribute the collected notes all at once at the beginning of the semester. The unexpected result, however, was that at the end of the semester a number of students complained on their questionnaires that I was lecturing straight out of my lecture notes!

Ah, the ungratefulness! I was at first disturbed by this lack of appreciation but have since changed my position. The students had a point: I was indeed lecturing from my lecture notes. And research showed that my students were deriving little additional benefit from hearing me lecture if they had read my notes beforehand. Had I lectured not on physics but, say, on Shakespeare, I would certainly not spend the lectures reading plays to the students. Instead, I would ask the students to read the plays before coming to lecture and I would use the lecture periods to discuss the plays and deepen the students understanding of and appreciation for Shakespeare.

In the years following the eye-opening experience described at the beginning of this paper, I explored new approaches to teaching introductory physics. In particular, I was looking for ways to focus attention on the underlying concepts without sacrificing the students' ability to solve problems. The result is *Peer Instruction*, an effective method that teaches the conceptual underpinnings in introductory physics and leads to better student performance on conventional problems. Interestingly, I have found this new approach also makes teaching easier and more rewarding.

The improvements I have achieved with *Peer Instruction* require textbook and lectures to play roles different from those they play in a conventional course. Preclass reading assignments from the book first introduce the material. Next, lectures elaborate on the reading, address potential difficulties, deepen understanding, build confidence, and add additional examples. Finally, the book serves as a reference and as study guide.

The ConcepTest

The basic goals of *Peer Instruction* are to exploit student interaction during lectures and focus students' attention on underlying concepts. Instead of presenting the level of detail covered in the textbook and/or lecture notes, lectures consist of a number of short presentations on key points, each followed by a *ConcepTest* —short conceptual questions on the subject being discussed. The students are first given time to formulate answers and then asked to discuss their answers with each other. This process *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.

Each *ConcepTest* has the following general format:

1.	Question posed	1 minute
2.	Students given time to think	1 minute
3.	Students record individual answers (optional)	
4.	Students convince their neighbors (peer instruction)	1–2 minutes
5.	Students record revised answers (optional)	
6.	Feedback to teacher: Tally of answers	
7.	Explanation of correct answer	2+ minutes

If most students choose the correct answer to the *ConcepTest*, the lecture proceeds to the next topic. If the percentage of correct answers is too low (say less than 90%), the teacher slows down, lectures in more detail on the same subject, and re-assesses 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.

Let's consider a specific example: Archimedes' principle. I first lecture for 7-10 minutes on the subject — emphasizing the concepts and the ideas behind the proof while avoiding equations and derivations. This short lecture period could include a demonstration (the Cartesian diver, for instance). Then, before going on to the next topic (Pascal's principle, perhaps), I show on the overhead projector the question shown in Fig. 4.

I read the question to the students, making sure there are no misunderstandings about it. Next, I tell them they have one minute to select an answer — more time would allow them to fall back onto equations rather than think. Because I want students to answer individually, I do not allow them to talk to one another; I make sure it is dead silent in the classroom. After about a minute, I ask the students first to record their answer and then to try to convince a neighbor of the rightness of that answer. I always participate with a few groups of students in the animated discussions that follow. Doing so allows me to assess the mistakes being made and to hear how students who have the right answer explain their reasoning. After giving the students a minute or so to discuss the question, I ask them to record a revised answer. As the results in Fig. 5 show, there is an overwhelming majority of correct answers after discussion. I therefore spent only a few minutes explaining the correct answer before going on to the next topic.

The convince-your-neighbors discussions systematically increase both the percentage of correct answers and the confidence of the students. The improvement is usually largest when the initial percentage of correct answers is around 50%. If this percentage is much higher, there is little room for improvement; if it is much lower, there are too few students in the audience to convince others of the correct answer. This finding is illustrated in Fig. 6, which shows the improvements in correct responses and confidence for all questions given during a semester. Notice that all points lie above a line of slope 1 (for points below that line, the percentage of correct responses after discussion is decreased). I consider an initial percentage of correct responses in the 40 to 80% range optimal and in subsequent semesters usually modify or eliminate questions that fall outside this range.

Figure 7 shows how students revised their answers in the discussion of the buoyancy question. As can be seen, 29% correctly revised their initially incorrect answer, while only 3% changed from correct to incorrect. Figure 6 shows there is always an increase and never a decrease in the percentage of correct answers. The reason for this is that it is much easier to change the mind of someone who is wrong than it is to change the mind of someone who has selected the right answer for the right reasons. The improvement in confidence is also no surprise. Students who initially are right but not very confident become more confident when it appears that neighbors have chosen the same answer or when their confidence is reinforced by reasoning that leads to the right answer.

It seems that sometimes students are able to explain concepts to one another more effectively than are their teachers. A likely explanation is that students who understand the concept when the question is posed have only recently mastered the idea and are still aware of the difficulties one has in grasping the concept. Consequently they know precisely what to emphasize in their explanation. Similarly, many seasoned lecturers know that their first presentation of a new course is often their best, marked by a clarity and freshness often lacking in later, more polished versions. The underlying reason is the same: as time passes and one is continuously exposed to the material, the conceptual difficulties seem to disappear and therefore become harder to address.

In this new lecturing format, the *ConcepTests* take about one third of each lecture period, leaving less time for straight lecturing. One therefore has two choices: (*a*) discuss in lecture only part of the material to be covered over the span of the semester or (*b*) reduce the number of topics covered during the semester. While in some cases (*b*) may be the preferable choice, I have opted for (*a*): I do not cover in class all the material covered in the text and in the lecture notes I pass out at the beginning of the term. I start by eliminating from my lectures worked examples and nearly all derivations. To make up for the omission of these mechanical details, I require the students to read the textbook and my lecture notes before coming to class. While this may sound surprising for a science course, students are accustomed to reading assignments in many other courses. In this way, students are exposed, over the length of the course, to the same amount of material taught in the conventional course.

Results

The results I have obtained — and which are supported by findings from other institutions where *Peer Instruction* has been implemented [6–8] — are striking. The advantages of *Peer Instruction* are numerous. The convince-your-neighbors discussions break the unavoidable monotony of passive lecturing, and, more important, the students do not merely assimilate the material presented to them; they must think for themselves and put their thoughts into words.

To assess my students' learning, I have used, since 1990, two diagnostic tests, the *Force Concept Inventory* and the *Mechanics Baseline Test*. [9–10] The results of this assessment are shown in Figs. 8 and 9 and in Table 1. Figure 8 shows the dramatic gain in student performance obtained on the *Force Concept Inventory* when I first implemented *Peer Instruction* in 1991. As Table 1 shows, this gain was reproduced in subsequent years. [7] Notice also how, in the post-test in Fig. 8, the scores are strongly shifted toward full marks (29 out of 29) and that only 4% of the students remain below the cutoff identified by Hestenes as the threshold for the understanding of Newtonian mechanics. With the conventional approach (Fig. 9) the gain was only half as large, in agreement with what has been found at other institutions for conventionally taught courses.

Do problem-solving skills suffer?

While the improvement in conceptual understanding is undeniable, one might question how effective the new approach is in teaching the problem-solving skills required on conventional examinations. After all, the restructuring of the lecture and its emphasis on conceptual material are achieved at the expense of lecture time devoted to problem solving. Development of problem-solving skills is left to homework assignments and discussion sections.

A partial answer to this question can be obtained by looking at the scores for the *Mechanics Baseline* test, which involves some quantitative problem solving. Table 1 shows that the average score on this test increased from 67% to 72% the year *Peer Instruction* was first used and rose to 73% and 76% in subsequent years.

For an unambiguous comparison of problem-solving skills on conventional examinations with and without *Peer Instruction*, I gave my 1985 final examination in 1991. Figure 10 shows the distributions of final examination scores for the two years. Given the students' improvement in conceptual understanding, I would have been satisfied if the distributions were the same. Instead, there is a marked improvement in the mean, as well as a higher cut-off in the low-end tail. Apparently, and perhaps not surprisingly, a better understanding of the underlying concepts leads to improved performance on conventional problems.

Feedback

One of the great advantages of *Peer Instruction* is that it can provide, in the *ConcepTest* answers, immediate feedback on student understanding. The tallying of 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 method. It gives a feel for the level of the class' understanding and allows the teacher to pace the lecture accordingly. The main drawback is a loss of accuracy, in part because some students may hesitate to raise their hands and in part because of the difficulty in estimating the distribution. A nice work-around is the use of so-called 'flash cards' — each student has a set of six or more cards labeled A–F to signal the answer to a question. [9] Other shortcomings are the lack of a permanent

record (unless one keeps data in class) and the lack of any data collected before the convince-your-neighbors discussion (a show of hands before the discussion influences the outcome).

2. *Scanning forms*. Because I am interested in quantifying the effectiveness of convince-your-neighbors discussion in both the short and the long term, I have made extensive use of forms that I scanned after class. On these forms, the students mark their answers and their confidence level, both before and after discussion. 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. For this reason, when using scanning forms, I always also ask for a show of hands after the question is answered for the second time.

3. *Handheld devices*. Since 1993 I have used a variety of interactive classroom responses systems (also known as 'clickers'). These systems allow students to enter their answers to the *ConcepTests*, as well as their confidence levels, on hand-held devices. Their responses are relayed to the teacher on a computer screen and can be projected so the students see it, too. The main advantage of the system is that analysis of the results is available immediately. In addition, student information (such as name and seat location) can be made available to the instructor, making large classes more personal; and some systems can also handle numerical and non-multiple-choice questions. Potential drawbacks are that the system requires a certain amount of capital investment and that it adds complexity to the lecture.

It is important to note that the success of *Peer Instruction* is *independent* of feedback method and therefore independent of financial or technological resources.

Conclusion

Using the lecture format described above, it is possible, with relatively little effort and no capital investment, to greatly improve student performance in introductory science courses — to double their gain in understanding as measured by the *Force Concept Inventory* test and improve their performance on conventional examinations. Despite the reduced time devoted to problem solving, the results convincingly show that conceptual understanding enhances student performance on conventional examinations. Similar benefits have been obtained in a variety of academic settings with vastly different student bodies. [6–8, 10] Finally, student surveys show that student satisfaction — an important indicator of student success — is increased as well.

Figure Captions

Figure 1: Conceptual (#1) and conventional question (#5) on the subject of dc circuits. These questions were given on a written examination in 1991.

Figure 2: Test scores for the problems shown in Fig. 1. For the conceptual problem, each part was worth a maximum of 2 points.

Figure 3: Correlation between conceptual and conventional problem scores from Fig. 2. The radius of each datapoint is a measure of the number of students represented by that point.

Figure 4: ConcepTest question on Archimedes' principle. For an incompressible fluid such as water, the second choice is correct.

Figure 5: Data analysis of responses to the buoyancy question of Fig. 4. The correct answer is A2. The pie charts show the overall distribution in confidence levels, and the shading in the bars correspond to the shadings defined in the pie charts.

Figure 6: (*a*) Percentage of correct answers after discussion versus percentage before discussion and (*b*) the same information weighted with the students'

confidence. Each data point corresponds to a single ConcepTest question. The filled data point is for the buoyancy question in Fig. 4.

Figure 7: How answers were revised after convince-your-neighbors discussion for the buoyancy question in Fig. 4.

Figure 8: *Force Concept Inventory* scores obtained in 1991 (*a*) on the first day of class and (*b*) after two months of instruction with the *Peer Instruction* method. The maximum score on the test is 29. The means of the distributions are 19.8 (out of 29) for (*a*) and 24.6 for (*b*).

Figure 9: (*b*) Force *Concept Inventory* scores obtained in 1990 after two months of conventional instruction. (*a*) For comparison, data obtained on the first day of class in 1991, 1992, and 1994. The means of the distributions are 19.8 out of 29 points for (*a*) and 22.3 for (*b*).

Figure 10: Final examination scores on identical final examination given (*a*) in 1985 (conventional course) and (*b*) in 1991 (*Peer Instruction*). The means of the distributions are 62.7 out of 100 for (*a*) and 69.4 for (*b*).

Table 1. Average scores for the *Force Concept Inventory* (FCI) and *Mechanics Baseline* (MB) tests before and after implementation of *Peer Instruction*.

		FCI					
Method of teaching	*Year*	*pre* [a]	*post* [b]	*gain*	*G* [c]	**MB**	*N* [d]
CONVENTIONAL	1990	(70%) [e]	78%	8%	0.25	67%	121
	1991	71%	85%	14%	0.49	72%	177
PEER INSTRUCTION	1993 [f]	70%	86%	16%	0.55	73%	158
	1994	70%	88%	18%	0.59	76%	216
	1995 [g]	67%	88%	21%	0.64	76%	181
	1996 [g]	67%	89%	22%	0.68	74%	153
	1997 [g]	67%	92%	25%	0.74	73%	117

[a] data obtained on first day of class; [b] data obtained after two months of instruction; [c] fraction of maximum possible gain realized; [d] number of data points; [e] no FCI pretest in 1990; 1991–1995 average shown; [f] no tests administered in 1992; [g] data obtained in 1995 and later reflects use of a new research-based text as pre-class reading.

References

[1] Sheila Tobias, *They're not dumb, they're different: stalking the second tier* (Research Corporation, Tucson, AZ 1990).

[2] Ibrahim Abou Halloun and David Hestenes, Am. J. Phys, **53**, 1043 (1985).

[3] Ibrahim Abou Halloun and David Hestenes, Am. J. Phys, **53**, 1056 (1985).

[4] Ibrahim Abou Halloun and David Hestenes, Am. J. Phys,**55**, 455 (1987).

[5] Hestenes, David, Am. J. Phys, **55**, 440 (1987).

[6] Sheila Tobias, *Revitalizing Undergraduate Science Education: Why Some Things Work and Most Don't*, Research Corporation: Tucson, AZ (1992).

[7] Catherine H. Crouch and Eric Mazur, *Am. J. Phys.*, **69**, 970-977 (2001).

[8] Adam P. Fagen, Catherine H. Crouch and Eric Mazur, *Phys. Teach.*, **40**, 206-209 (2002).

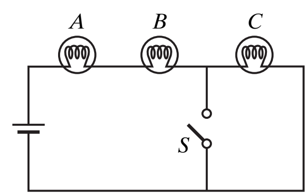
[9] D. Hestenes, M. Wells, and G. Swackhamer, Phys. Teach. **30**, 141 (1992).

[10] D. Hestenes, M. Wells, and G. Swackhamer, Phys. Teach. **30**, 159 (1992).

[11] David. E. Meltzer and Kandiah Manivannan, Phys. Teacher **34**, 72 (1996).

[12] R.R. Hake, AAPT Announcer **24** (1), 55 (1994).

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 stay 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 circuit shown, calculate (a) the current in the 2-Ω resistor and (b) the potential difference between points P and Q.

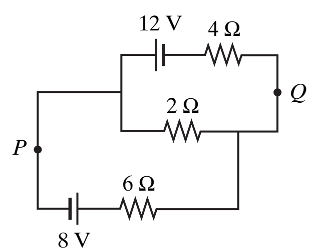


Figure 1 — Eric Mazur

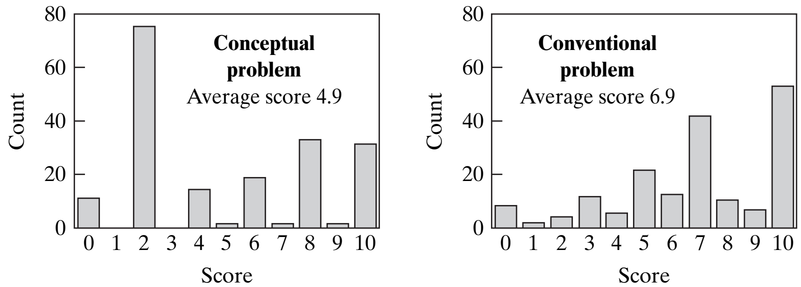


Figure 2 — Eric Mazur

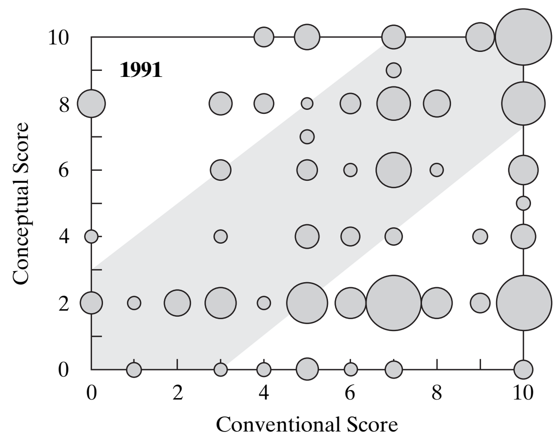
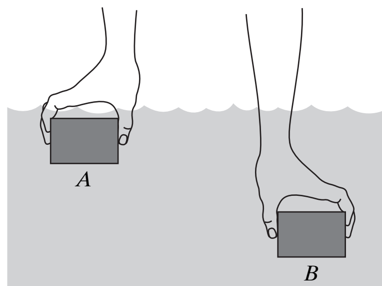


Figure 3 — Eric Mazur

BUOYANCY

Imagine holding two bricks under water. Brick *A* is just beneath the surface of the water, while brick *B* is at a greater depth. The force needed to hold brick *B* in place is



1. larger than
2. the same as
3. smaller than

the force required to hold brick *A* in place.

Figure 4 — Eric Mazur

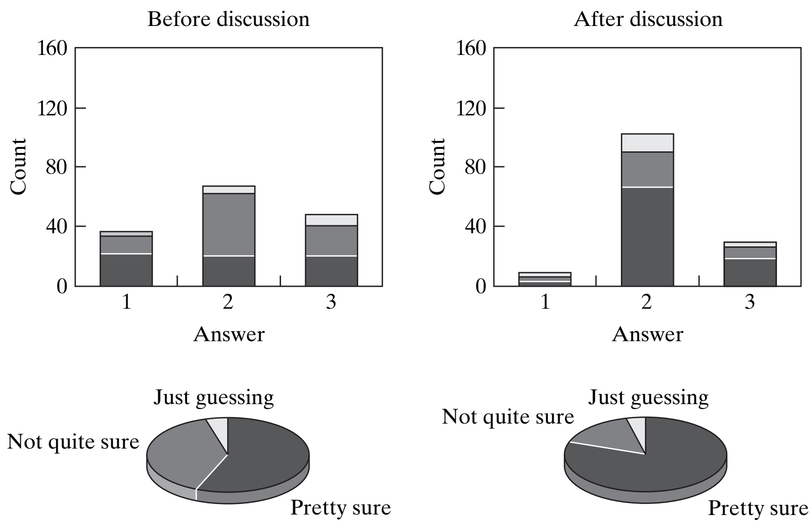


Figure 5 — Eric Mazur

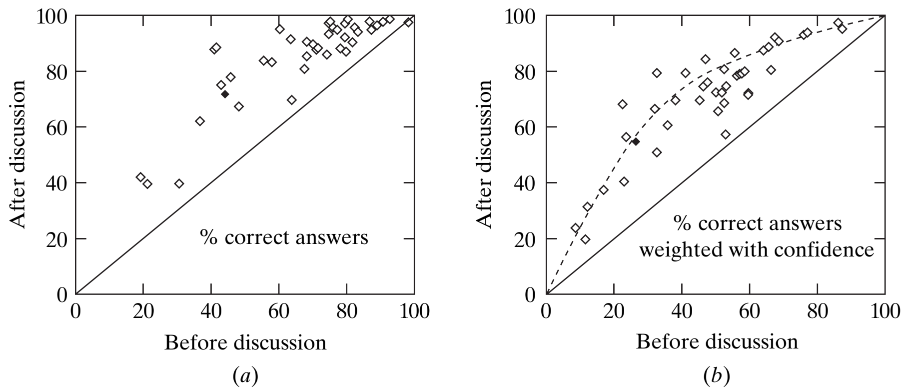


Figure 6 — Eric Mazur

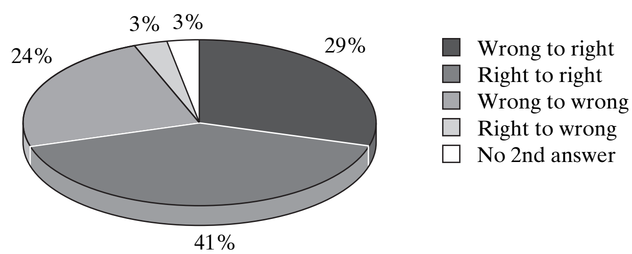
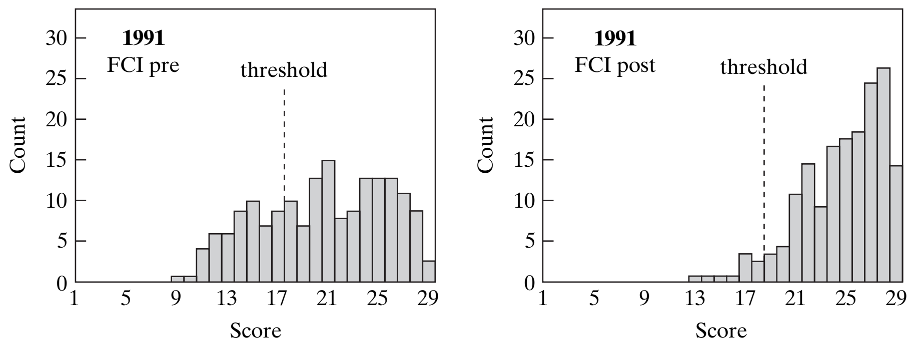
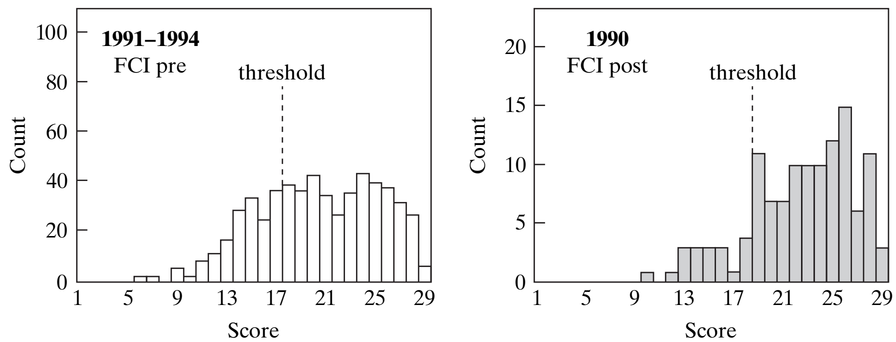


Figure 7 — Eric Mazur



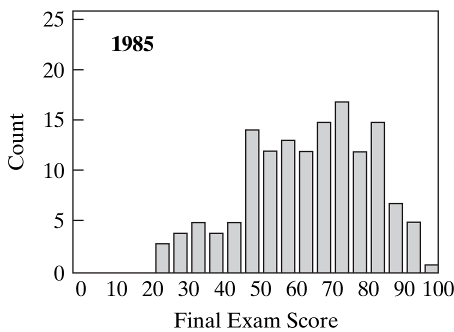
(*a*) (*b*)

Figure 8 — Eric Mazur

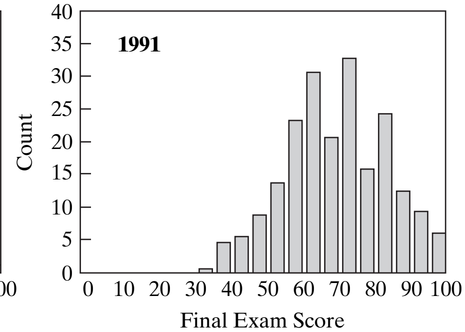


(*a*) (*b*)

Figure 9 — Eric Mazur



(*a*)



(*b*)

Figure 10 — Eric Mazur