Reinventing college physics for biologists: Explicating an epistemological curriculum

Edward F. Redish and David Hammer
Departments of Physics and Curriculum and Instruction, University of Maryland, College Park, Maryland 20742-4111

(Received 29 July 2008; accepted 25 March 2009)

The University of Maryland Physics Education Research Group has done a five-year project to rethink, observe, and reform introductory algebra-based (college) physics, which primarily serves life-science majors. We refocused the class on helping the students learn to think scientifically—to build coherence, think in terms of mechanisms, and to follow the implications of assumptions. We designed the course to tap into students' productive conceptual and epistemological resources, based on a theoretical framework from research on learning. The reformed class retains its traditional structure in terms of time and instructional personnel, but we modified existing best-practices curricular materials. We provided class-controlled spaces for student collaboration, which allowed us to observe and record students learning directly. We also scanned all written homework and examinations and administered pre-post conceptual and epistemological surveys. The reformed class enhanced the strong gains on pre-post conceptual tests produced by the best-practices materials while obtaining unprecedented pre-post gains on epistemological surveys instead of the traditional losses.

[DOI: 10.1119/1.3119150]

I. RETHINKING ALGEBRA-BASED PHYSICS FOR BIOLOGY MAJORS

Algebra-based (college) physics is one of the largest service courses in most physics departments. At the University of Maryland we teach approximately 800 students a year in each term of this two-semester class. The population is increasingly dominated by science majors, including pre-health care, such as premedical, predental, pre-physical therapy, preveterinary, and a growing number of pre-research biologists.

In the years 2000–2005, the University of Maryland's Physics Education Research Group carried out an NSF-supported research study to observe student behavior in algebra-based physics and to explore reforms in the course.[1] The reforms we created for the class were based on our reading of the current needs of modern biology students,[2–4] interviews with biology faculty, a theoretical framework that gives us insight into how students think and learn about physics,[5–7] and our experiences in small seminar courses for college students and in high school courses.[8–10]

Every course contains not only explicit content but elements that are traditionally not made explicit in descriptions of the course—an implicit curriculum.[11] For example, traditional instructors tend to assume that students learn how to think about and do scientific reasoning while doing traditional class activities, such as reading the text and doing end-of-chapter problems. Some students do learn how to think scientifically successfully, but research indicates that most do not and some pick up bad habits and inappropriate modes of thinking.[12] We chose to focus the course on helping students learn how to learn science, content that is implicit in most courses and that research has convinced us needs to be addressed explicitly.

Many of these implicit elements are epistemological issues about the nature of scientific knowledge: how we know what we know, how to create new knowledge by problem solving, how we make inferences, what makes sense, and how to build physical intuition. These issues have particular importance for the biology majors currently dominating college physics classes, but they are equally important for other populations of physics students. The implicit epistemological content tacitly taught in traditional courses often is not what we want our students to learn; rather, it encourages poor approaches to learning such as rote memorization and the denigration of everyday experiences and intuitions.[13]

Students bring epistemological assumptions into our classes as a result of their experiences, especially in their previous science classes. To understand the match or mismatch between student epistemologies and what we want them to learn, we transformed the course to encourage student learning to take place in class-managed areas where it could be observed and videotaped. We collected large amounts of written data, including pre-post conceptual and epistemological surveys, and digital scans of all written homework and exams. The instructor encouraged students to reflect briefly on the class in written essays. In addition, researchers in our group who were not part of the instructional team interviewed some students about their experiences before and after the class. In this paper, we present the reforms we developed and review the evidence of their success.

We achieved what we believe to be the first documented large gains on an epistemological survey in a large lecture introductory physics class. We did it while not only retaining but enhancing high values for the fractional gain on a mechanics conceptual survey. We produced large gains compared to traditional classes on a split task postinstruction concept survey that measured not only students' knowledge of the correct results but their intuitive comfort with those results. We also documented the kinds of epistemological difficulties students encounter during the course and the extent to which those difficulties can be overcome. These observations were done within the context of a traditional environment with the same resources provided to our standard large lecture class.

In Sec. II, we describe our motivations for choosing the reforms we did. In Sec. III, we describe the theoretical basis for our analysis of our goals and the instructional tools we chose to reform. We describe the reforms we carried out in Sec. IV. In Sec. V, we describe our methods for observing and evaluating the class, and we present our observations and conclusions in Sec. VI. What we learned from detailed research studies that contribute to our understanding of how an individual student learns physics is described in Refs. 14–18.

II. DECIDING WHAT MATTERS

One of the most significant transformations in science in the past half-century has been the growing strength of biology as a fundamental science. There is broad agreement among leading biology and medical researchers that future biology students will need to become much more knowledgeable in basic physics, chemistry, and mathematics. These students require not only a familiarity of the facts and vocabulary of those fields but a deep understanding of the disciplinary patterns of knowledge and process, including a solid understanding of scientific reasoning. Hence, it is essential that physics education go beyond isolated facts and narrow procedures. More than helping students understand established ideas, science instruction must help them understand how those ideas came to be. Students must be prepared to contend with ambiguities, make sound judgments about what to accept and what to question, reconsider past assumptions, and adapt to new discoveries. They must learn what a measurement means and does not mean. They must learn how to evaluate their data and see its implications. In short, they must learn an adaptive expertise—the ability to respond effectively and productively to new situations and new knowledge as it develops.[19]

Science instruction at the university level tends to ignore an explicit focus on helping students develop these elements of adaptive expertise, hoping that they will spontaneously spring into being as a “side effect” of traditional coverage of traditional content. This traditional approach works for a small minority of students after many years of combined undergraduate and graduate training. Our goal in this project was to learn how to help more students develop these broad thinking and learning skills by paying explicit attention to these issues and by developing a curriculum to deal with them.

A. A resource-based model of mind

Our redesign was based largely on a resources based view of students’ knowledge and reasoning[5–7] that supports Einstein’s claim that “The whole of science is nothing more than a refinement of everyday thinking.”[20] Everyday thinking involves both conceptual and epistemological resources, and learning physics begins by marshalling those resources in productive ways.

Student conceptual resources include their extensive intuitive knowledge about physical phenomena and causal mechanisms,[21] everything from what would happen if someone tried to kick a bowling ball to what it feels like underwater, from how an oven mitt can keep them from getting burned to how a source of light or odor feels stronger up close than far away. Students use a rich but highly fragmented variety of knowledge and experience as they interact with the physical world. Students should draw on those resources while reasoning about questions in physics. In many cases, the ways they are inclined to draw on those resources lead them to wrong conclusions. Students’ reasoning that current is used up in light bulbs, for example, draws on resources that would be productive for thinking about how fuel is used up in gas lanterns. The solution for students thinking of current being used up is not for them to stop using their common sense. It is for them to find other aspects of common sense to apply, other resources in their repertoire, such as those they would use to understand how trying to hold a moving rope can burn their hand. Rather than set their common sense aside, students should search within it for other possible conceptual anchors.[22]

A resource-based model of conceptual knowledge takes a dynamic view of thinking that is in apparent contrast to accounts of novice understanding in terms of coherent “naïve theories” and misconceptions.[23] Research on the latter has established patterns of student reasoning that differ from expert understanding, and these findings have been interpreted to suggest that intuitive knowledge is an impediment to expertise. In some important respects that interpretation is the opposite of what the original research established,[24] which was that novice “misconceptions” represent sensible reasoning well-grounded in experience. Much of the difficulty is that the naïve-theories account views intuitive knowledge as unitary, seeing the misconceptions as the one way students have for thinking about the topic. Teachers and researchers who have close contact with students know that students have many ways of thinking. Common sense does not have a coherent organization; it is made up of many parts, and the common sense answer to a question depends on which parts are activated at a particular instant. A resource-based view provides an account of that variability and of how science can genuinely be a refinement of everyday thinking.

The core innovations of our reform attend explicitly to student epistemologies; that is, to how students understand knowledge and learning in physics. Just as students have a vast collection of resources for thinking about physical phenomena and mechanisms, they have a vast collection for thinking about knowledge, about its various forms and sources, and about how it can arise and be used in various sorts of activities. Just as they use their collection of conceptual resources for experiencing and making sense of the physical world, they use these epistemological resources for experiencing and making sense of knowledge and learning. Depending on the situation, they use different epistemological resources for thinking about what knowledge entails, the forms it takes, how it arises, and whether it is valid.

In traditional physics courses, students often learn to set their everyday experience aside.[25,26] They frame the task as a matter of receiving and rehearsing information, information that need not make sense. A primary goal in our courses is to help them frame learning in other ways, tapping productive epistemological resources for thinking about sense-making and argumentation, for understanding physics knowledge as a coherent system of ideas rather than a collection of independent pieces of information. We pursue this goal both explicitly in the instructions and advice we give students and implicitly in the structure and design of assignments, lectures, tutorials, and labs.

III. TRANSFORMING THE CLASS STRUCTURE WITHIN EXISTING CONSTRAINTS

A. The traditional teaching environment

Because the perception of a reform depends on what it is being compared to, we describe briefly the traditional environment for algebra-based physics as it was at the University of Maryland when we began the project in 2000.[27] Our traditional algebra-based physics class is taught in two fourteen-week semesters covering the topics of "mechanics, heat, sound, electricity, magnetism, optics, and modern physics."[28] Each half of the class is taught to 400–500 students per semester, divided into three lecture sections of 100–200 students each. Each lecture section is assigned to a faculty member who is responsible for the content, lectures, assigning reading, and homework. Each lecture section is divided into small group sections of 24 students. Each small group section meets for one 3-hour period per week run by a graduate teaching assistant (TA). The first hour of the period is typically a problem solving recitation; the last two hours are a laboratory. The students purchase a common text, which is typically the source of all homework problems, and a laboratory manual.

Each professor makes a somewhat independent choice as to what content to emphasize within the constraints of the catalog description. Although there is some variation, an attempt is made to keep the first semester fairly common because a significant fraction of students switch from one lecture section to another after the first term. Homework is handled idiosyncratically. Homework may be assigned from the book or from an on-line homework system and may or may not be graded. Laboratories are traditional with extensive write-ups and step-by-step guidance provided. Students work in pairs and create individual lab reports. Ten laboratories are required each term and makeup periods are provided during two weeks of the term in which students can complete missed labs. A separate faculty member is responsible for the laboratories and for training the TAs in managing the lab.

The lecture faculty are responsible for creating, grading, and managing the examinations for their own students. There are no common exams. Typically, there are two to three midsemester exams and a final. Sometimes exams are multiple choice or short answer, but they often include problems and require calculations. Recitation-section TAs are typically recruited to do much of the grading. Typically, the only interaction between the lecture and the lab part of the class is that the same TAs run the lab and recitation sections.

Traditional lectures include demonstrations, derivations, and sample solutions to homeworklike problems. There is rarely much interaction with the students during lecture. Attendance during lecture varies from instructor to instructor and ranges from 25 to 85% of the registered students. Typical recitations are run by TAs as problem-solving mini-lectures, with the choice of problem sometimes guided by student questions. If the recitation does not contain a required quiz and if the TA has been instructed not to solve the current week's assigned problems, the attendance is typically less than a third of the registered students.

One of us (EFR) taught algebra-based physics in this traditional mode for many years reasonably successfully, meaning there was good attendance in lecture (typically more than 75%), high ratings in end-of-year evaluations from students (above departmental averages for the class), and some anecdotal successes (individual students reporting relief and delight that the course was not as impossible as they expected).

B. The reformed teaching environment

During the five years of the project, the authors were the lecturers of record for a semester of the course 11 times.[29] As a result of our reconsideration of the course goals and on the basis of our resource framework of student learning, we reformed each of the components of the class to be explicit about epistemology. There are many research-based reforms that help build students' conceptual knowledge. Many of these reforms are based on a cognitive-conflict[30] or elicit-confront-resolve[31] pedagogical model in which students are asked to make predictions so as to display their intuitions. The students then see empirical results that show these intuitions are incorrect, and finally the instructor helps the students resolve the conflict. Our experience as teachers and researchers has been that this model often has the negative epistemological side effect that students learn to consider their intuitive knowledge and experience as irrelevant for physics learning; they learn to set it aside, rather than to draw on and refine it. To avoid this effect, we modified each of these conceptually oriented reforms.

1. The lecture

We implemented three reforms that increased the epistemological emphasis of the lectures: explicit epistemological discussions, adaptations of the Peer Instruction materials,[32] and the use of epistemologically modified *Interactive Lecture Demonstrations*.[33]

Being explicit about epistemology

We made the epistemological framing of the course explicit, including through the use of a vocabulary we introduced early in the semester. We designed this vocabulary based on previous work in a small seminar class.[10] One of us (EFR) created and used a series of icons to use in PowerPoint slides and course materials to help reinforce and remind students of the various epistemological framings.[34] The terms include: *shopping for ideas, sense making, seeking coherence, restricting the scope, choosing foothold ideas*, and, *playing the implications game*.

Shopping for ideas. The overarching message of the course is that "the whole of science is nothing more than a refinement of everyday thinking." Hence a core activity of the course needs to involve students becoming more familiar with, and critically aware of, their everyday thinking. We use the metaphor of "shopping" to help students think of their own knowledge and experience as having a large inventory of possibilities through which they can browse. We give an example of how to do this with a story to connect to everyday epistemology: Imagine you have met a new person and there's something about him that bothers you, but you can't put your finger on what it is. So you think about it, trying to figure out whether he reminds you of someone or you've met him before. You "shop" in your mind through different sections of your knowledge and experience. You ask "Have I met him before?" and try different possibilities: "Have I seen him at the pool? At the store? In art class?" "Who does he remind me of?" Eventually you may realize that he looks and sounds a bit like a character in a movie you saw recently.

Now you know that it's not really this new guy who troubles you but that movie character, and you don't have to worry about it any more. Or, if you were to realize that you've met him before and had an unpleasant interaction, you'd have found that feeling of irritation is warranted.[10]

This sort of shopping in their minds serves two purposes. One is to help students locate the origins of an impression they have about some problem ("I can't explain why, I just think that's what will happen.") as well as to help them locate alternative possibilities ("We're saying electricity flows; maybe we should think of other things that flow and try to compare.").

Sense making. Students in many college science classes have the view that science is a collection of unrelated facts that do not necessarily need to be comprehensible.[25,26] A dramatic example took place in a videotape of a lesson trying to help students build analogies for thinking about electric current. One student asked the TA to stop using analogies and tell them how current really worked. The TA responded, "What do you want me to do, give you a bunch of words that you don't know what they mean?" The student answered (with a straight face), "Well, that's what I'm used to." We emphasize to students that the principles, definitions, and equations of physics should make sense—that they should be able to restate principles, definitions, and equations in their own words and explain clearly what they are saying.

Seeking coherence/safety net. Physicists take for granted that knowledge in physics should cohere; we come to accept ideas and findings as true because of the way they hold together with other ideas and findings, ideally with all the other ideas and findings of which we are aware. When there are conflicts, we need to resolve them, and if we cannot, it tempers our confidence in the conclusions and our satisfaction with our understanding. Many students are more accustomed to thinking of physics knowledge as a set of facts and formulas, independent pieces of information to remember, and they have framed learning science as a matter of memorizing information.[10,16,35] We try to guide them to a more productive framing, looking for resources in their experience that might help.

We also stress to students that their "one-step" recall memory can be unreliable. The mind reconstructs memory from bits and pieces, and something remembered may crosslink to distinct memories.[36] Having coherence means there is numerous cross-linking in memory that provides them with a safety net that provides stability and consistency to their reasoning—a stability that is not present when they memorize independent results.

Restricting the scope. One of the challenges of learning physics is learning to ignore some of what happens in the real world in order to construct models. If students frame learning physics as learning about the real world all at once, they will constantly be frustrated and confused by the routine practices in physics of making simplifying assumptions, positing idealized conditions, and ignoring some aspects of the physical world. We make this restriction of what is being considered an explicit topic of discussion, how cordoning off a portion of the world can serve as a step toward understanding the world more generally. We make a point of noting explicitly when cordoning off a portion of the world is taking place.

Choosing foothold ideas. Students often find themselves in the position of not knowing what to believe; that position is, of course, common to science experts as well. We introduce the notion of a foothold idea as one we choose to accept as true, at least for the time being, as a way to proceed. As we find other ideas and findings fit with a foothold idea, and as we are able to respond to counterarguments and counterevidence, we form a greater and greater commitment to the foothold; we are willing to work harder to reconcile other reasoning to fit with it. For example, if an experiment produces evidence of fusion occuring at low temperatures, in contradiction to strongly held foothold ideas about nuclear and atomic physics, or if measurements show that the expansion of the universe is accelerating in contradiction to foothold ideas about the make up of the physical universe, we maintain skepticism of the results and work hard to discredit them to reconcile the contradiction in favor of the footholds. Sometimes, it becomes too difficult to reconcile the contradictions with current foothold ideas, and we choose new ones.

Playing the implications game. Having chosen a foothold idea, we consider its implications; if X is true, what would that mean? Often that leads us to something we can't accept, and we abandon X. Sometimes it leads to surprises that turn out to be true. Students might not apply this reasoning to learning physics without prompting if they frame what they are doing as only remembering information. We identify the "implications game" to let students know that is what we're doing.

Peer Instruction

Starting in 2002, we adapted elements of the *Peer Instruction* (PI) environment[32] for this class. Each student was issued a remote answering device (clicker). The instructor periodically asks a multiple-choice question during the lecture to which the students respond using these devices. A computer automatically displays a histogram of the results.

In the original PI environment a clicker question typically follows a 10–15 minute lecture segment. The student is asked to think about the answer individually and then choose an answer. If the question is well designed, the class should display a mix of answers. The students are then given two minutes to discuss the question with their neighbors and again choose an answer. If there is now large-scale agreement among the students on the correct answer, the lecturer goes on. If not, the lecturer adds a brief lecture segment to explain the correct answer.

We modified this reform as follows. In all cases, the instructor draws on the class for discussion of the question; sometimes the instructor presents the question alone and asks the class to suggest possible multiple-choice answers. In part to help save students from embarrassment and in part to encourage the habit of mind, we ask students to generate answers and reasoning that they think someone who had not studied physics might believe. Discussion often focuses on students' intuitions based on their real-world experience. After the first click, students have the opportunity to defend or challenge answers (not necessarily their own). Once the correct answer is known, further discussion focuses on the wrong answers, why they were chosen, and whether they had a "correct" intuitive core. The goal is to encourage students to not just "know" the right answers, but to understand them as both plausible and intuitive. One of us [DH] often creates spontaneous clicker questions "on the fly" in response to a student question or to a sticky point in the lecture. The other

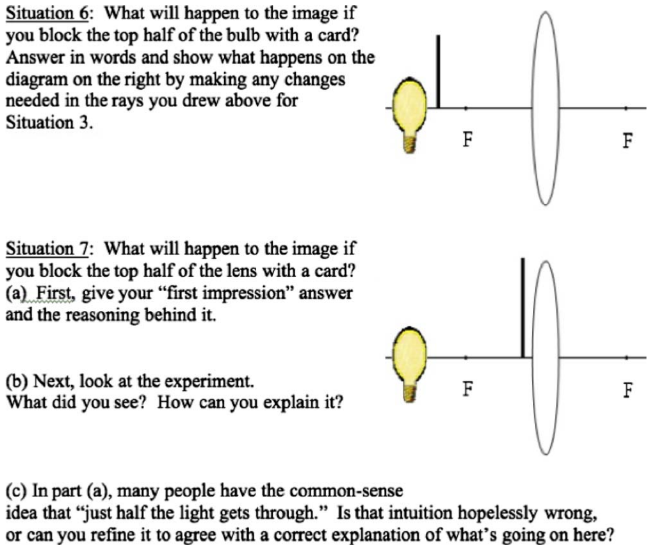
[EFR] tends to have pre-prepared questions and to include them (without answers or discussions) in his PowerPoint slides, handouts of which are distributed via the web the night before the class. See some sample problems in Ref. 37, Figs. S3 and S4.

Interactive lecture demonstration

We increased the interactivity of our lectures by adapting the Interactive Lecture Demonstration (ILD) environment.[33] In its original form, a full lecture is devoted to a set of connected demonstrations. The choice of topics relies heavily on education research to determine critical areas where students tend to show or develop misconceptions that interfere with their understanding of the physics being presented. The method relies heavily on cognitive conflict, with an elicit-confront-resolve instructional model. Students receive two identical worksheets, one for their predictions, made after the experiment has been explained but before it is done, and one for the results they observe in the experiment. At the end of the demonstration, the students hand in their predictions. They get a small amount of class credit for having participated in the ILD. They keep their results sheet.

We liked the ILD model for its interactivity, its proven success in developing conceptual knowledge, and its guided inquiry structure. We were concerned that the method might work against our epistemological goals. The cognitive conflict approach can send students the message that their intuitions about the physical world are generally misleading and irrelevant to a physics class. This message might contribute to what we have documented in students' epistemologies; that is, they learn to set their intuitive knowledge aside, rather than refine it.[25,26] The two-worksheet structure of ILDs embodies that view: Students hand in the page with their predictions (perhaps to show the instructor how wrong they were before the lesson) and they keep the page with the "right answers." Although we have no evidence that this kind of activity directly contributes to the kind of problems we have observed in other environments, we wanted our course to send a consistent metamessage to students about their intuitions and how to use them.

We therefore modified the ILD approach so that students receive only a single worksheet that emphasizes finding the valid content of a student's intuition and refining it. The lecturer guides students through the worksheet and leads a discussion about the issues it raises. We developed about a half-dozen worksheets to be used in each semester.[38] An example of an epistemologized discussion from an ILD worksheet is shown in Fig. 1. In this example, a discussion of the "misconception" that blocking half a lens will result in blocking half the image (instead of reducing the intensity but showing the full image)[39] is paired with discussion of blocking half the bulb, which does result in blocking half the image. By presenting the two situations, we hope to help students refine their existing intuitions and to avoid implying that their existing intuitions are systematically wrong. The students are not graded on their answers to either the clicker questions or the ILDs, but they are given points for doing them. We design homework and test questions to help them assess their understanding of the material discussed during ILDs.


Situation 6: What will happen to the image if you block the top half of the bulb with a card? Answer in words and show what happens on the diagram on the right by making any changes needed in the rays you drew above for Situation 3.

Situation 7: What will happen to the image if you block the top half of the lens with a card?
(a) First, give your "first impression" answer and the reasoning behind it.

(b) Next, look at the experiment.
What did you see? How can you explain it?

(c) In part (a), many people have the common-sense idea that "just half the light gets through." Is that intuition hopelessly wrong, or can you refine it to agree with a correct explanation of what's going on here?

Fig. 1. A component of an epistemologized ILD worksheet. The worksheet is done in lecture. Students discuss the issues among themselves; the instructor leads the discussion and shows the demonstrations.

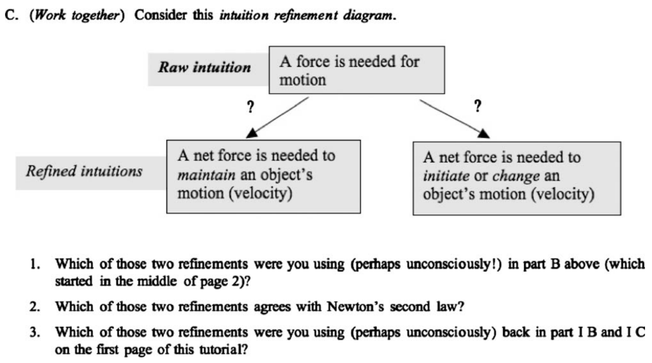
2. *The recitation*

In our experience with large lecture introductory physics classes over many years, we have observed many TA-led problem-solving recitation sections. We found that the solutions developed and presented by the TAs sometimes undermined the approaches we were trying to foster in lecture and occasionally contained conceptual or mathematical errors. The students rarely engaged in questioning the TA's presentation in the depth that would have made it useful to them. Current research in our group supports these informal impressions.[40] Unless we required attendance at recitation by giving a graded quiz, we found that the number of students attending recitation typically dropped to 1/3 of the registered students or fewer. We therefore did not hesitate to replace the problem solving recitation with a conceptual tutorial. Other mechanisms were created to provide students a venue to have questions answered about the homework.

3. *Tutorial*

Instead of a TA-led recitation, students work through worksheet-based group-learning activities based on the model developed at the University of Washington (UW).[41] We began with that set of tutorials and the Activity-Based Physics Tutorials.[42] These tutorials were designed to produce conceptual gains and have been demonstrated to succeed in this goal,[43] but our research indicated that they did not help develop better epistemological attitudes associated with the class.[26] We conjectured that part of the problem is that the cognitive-conflict approach stresses the failure of everyday intuition. So we created some new tutorials, again with epistemologies in mind.

Our "epistemologized" tutorials emphasize the reconciliation of everyday, intuitive thinking and experience with formal scientific thinking and encourage explicit epistemological discussions about the learning process. A common tool that we employed was the paired-question technique developed by Elby.[9] Instead of offering students introductory questions that research has shown most will answer incorrectly, a pair of matched questions are created so that most students will get one right and one wrong.[44] Elaboration and analysis of the pair of answers shows students that their intuitions are leading them into a conflict. They are then



C. (*Work together*) Consider this *intuition refinement diagram*.

1. Which of those two refinements were you using (perhaps unconsciously!) in part B above (which started in the middle of page 2)?
2. Which of those two refinements agrees with Newton's second law?
3. Which of those two refinements were you using (perhaps unconsciously) back in part I B and I C on the first page of this tutorial?

Fig. 2. A section from a tutorial worksheet containing a reconciliation diagram and an epistemological discussion. In the actual worksheet, space is left for the students to write their answers.

guided to use the (essentially correct) "raw intuition" that underlies both their answers and are guided to maintain that intuition and to refine it. In this way, we hope to convince them that their intuitions about the physical world are valuable and, when properly refined, support their new physics knowledge. An example of a "reconciliation diagram" from a tutorial on Newton's second law (unbalanced force goes with acceleration, not velocity) is shown in Fig. 2.

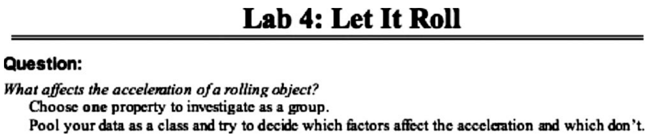
In other ways the sections proceed in the usual UW Tutorial fashion, with two trained facilitators wandering the room, listening, asking questions, and checking results.

4. Scientific community labs

Although our original plan did not call for reforming the laboratory, watching videotapes of students responding in lab changed our minds.[45] We observed students "going through the motions" in following the explicit protocols given in the lab manual. They spent little or no time trying to make sense of what was happening or trying to relate the procedures or the results to the physics they were learning in the class.[46,47] Students even made comments to the effect that they did not expect to make sense of what was happening, which some students found distressing. In this way, the laboratories sent messages about the nature of physics knowledge and how it was acquired that contradicted the messages we were trying to send. We therefore spent time developing and refining the laboratories. The result was the scientific community labs.[48,49]

The reformed labs are held for periods of two hours with 20–24 students and one TA. The goal is to help students understand the construction of knowledge through measurement and analysis. Instead of an already-set-up apparatus and detailed lab manual, we give students a half-page instruction sheet containing a one-sentence question that can be answered empirically, such as "What affects the acceleration of a rolling object?" or "How does the force between two magnets change if you change the distance between them?" The students' task is to design an experiment using the available equipment, make measurements, analyze the results, and present them to the class. We choose questions that make designing and carrying out the experiment feasible.

Unlike traditional practices, we do not use labs as ways to follow up on the theoretical discussion in lecture. For example, the lab shown in Fig. 3 takes place well before rotational motion is discussed in lecture. In this way, the purpose of the lab is centered on students' understanding of physics as an empirical science.[48] The labs can introduce students to a topic phenomenologically or prepare them for later theoretical development, as is often the order of things in real scientific research and discovery.



Lab 4: Let It Roll

Question:

What affects the acceleration of a rolling object?

Choose one property to investigate as a group.

Pool your data as a class and try to decide which factors affect the acceleration and which don't.

Fig. 3. A typical laboratory handout. We give students suggestions for how much time to spend on the various components of the experiment, but we do not give any other suggestions as to what to do.

Students work in groups of four, write group reports, and are evaluated on their thoughtfulness, persuasiveness, and understanding of measurement concepts and on the clarity of their discussion of how they could improve their experiment if they were to repeat it. During the lab session, there is considerable interaction and discussion among the students and with the TA. We give students two weeks per laboratory, which gives students four hours to plan, implement, analyze, and discuss each experiment.

There is a practical benefit for TAs; students collaborate on reports in groups of four, and they hand in reports every other meeting, which reduces the number of lab reports a TA needs to read and grade by a factor of eight. These readings can be more thorough, and TAs have time to give detailed feedback.

5. Homework

Traditional algebra-based physics classes typically assign many end-of-chapter problems for homework. Many are exercises that focus on the process of manipulating equations and numerical answers. Other more substantial problems may also be assigned. Early in the project, we noted informally that many students concentrated their effort on the exercises, which they could do without much thought or understanding, but they would give up on the more difficult problems before getting very far. That is, we saw our students behaving in ways Schoenfeld observed in mathematics classes—students seemed to believe they should be able to solve "any assigned problem in five minutes or less."[50]

We decided to drop all exercises and instead designed homework assignments entirely around challenging problems that require students to think and make sense of the ideas. Students have experience with problem sets before our class, and many still attempt to use "exercise methods" on more challenging problems,[16] not expecting they would have to spend much time working on problems outside of class for an introductory course. For this reason, we emphasize early and often in the course that they should expect to spend anywhere from 15 to 60 minutes per problem, ideally working in groups, discussing the issues with each other.

In accordance with this expectation, we assign only about five problems each week. They include a mix of challenging activities including representation translation problems, context-based reasoning problems, ranking tasks, estimation problems, and essay questions with epistemological content.[51] Examples of the first two types of questions are available,[37] and a collection that includes many of the problems we have developed is also available.[52] An independent short "Tutorial Homework" assignment is also given in conjunction with the Tutorial work. This assignment is a brief review of tutorial ideas and is expected to take students about 15–30 minutes per week.

We have had to compromise substantially on the extent and quality of the feedback we can give students on their homework. In order not to overload the TAs, we choose one problem each week for careful grading on a five-point range with written feedback. TAs grade other problems without written feedback for what we describe as "honest effort." We provide elaborate solutions on the course website after the homework is due and enjoin students to study them; students need to learn that a good score on their assignment did not necessarily mean that they had done the problem correctly due to the "light" grading.[53]

Although this population of students has a reputation for being weak in mathematics, we did not find that the basic manipulations were something that needed to be practiced extensively for most of our students. Rather, the issue of making sense of equations and translating between equations and physical experience was difficult. That mathematical expressions can express ideas and reasoning was an explicit topic in lecture and was modeled with examples and our Peer Instruction problems. Detailed solutions were made available on the class website, including extensive discussions of motivation and interpretation.

6. The course center

Because we converted the traditional discussion sections to tutorials, they no longer provide opportunities for students to discuss the homework problems. To close this gap, we set up a "course center" staffed by TAs or the instructor approximately twenty hours per week, where students can gather to work on homework. The TA or instructor offers assistance and coaching on good problem solving strategies, not solutions. We found that it was necessary to arrange the furniture in the room to discourage the TAs from making presentations to the entire class. There is now no central writing space, and the tables are arranged so that the students sit facing each other. White boards are available, but only at places behind the tables where seated students can easily reach them but the TAs cannot.[54]

7. Exams and quizzes

We design exams to require the kind of thinking we want students to learn, and we deliver them in a way that communicates that they are to be used as formative rather than as purely summative evaluations. The exams and quizzes include items that call on the students to use the epistemological skills we are trying to help them develop. One of us (EFR) follows a strict structural pattern on hour exams and explains it to the students. Every exam contains five questions: A multiple-choice multiple-representation question (25 points), two long-answer problems (25 points each), an estimation problem (15 points), and an essay question (10 points). On the long-answer problems the answers are worth 5–10 points, and the explanations and reasoning are worth 15–20 points. On the estimation problems the answers are worth 3 points (and a wide range of answers are acceptable), and the method is worth 10–12 points. The estimation questions require the creation of numbers from one's personal experience, and the experience has to be explained to receive full credit. The reduced emphasis on getting an accurate answer and the added stress on reasoning is in response to the tendency of students in this population to focus only on answers and to ignore reasoning. See Ref. 37, Figs. (S3), (S5), and (S6), for a sample multiple-choice multiple-representation question, essay question, and estimation problem.

Exams were typically given in the last class of a week, graded over the weekend, and returned to the students in the first class of the next week. The exam and the grading pattern were then discussed in detail during that class. The discussion included not only the "right" answers but also a discussion of the common errors, misconceptions, and difficulties that many students encountered.

Makeup exams and regrading

Most of the students in the class are juniors and seniors with much experience in other science classes, with expectations about what we would put on a test. At the beginning of the project we often heard students say something like, "Science exams have so much time pressure that you have no time to think during an exam and the profs can't expect you to. So you have to memorize stuff so you can give it back quickly on an exam."

These expectations help explain why many students do poorly on the first exam in the course. For them the feedback is negative and distressing and occurs about 1/3 of the way into the semester, with some critical material already behind them. One option would be to give three hour-exams a term and drop the lowest grade, but this option could send the message that if they did poorly on the first exam, they could let that material go. To the contrary, we want them to understand that physics learning is highly cumulative; they need to think that is important to go back and learn the material if they have not understood it well.

An important goal of the exam structure is to help the students learn to use their exam results to focus on identifying problems in their thinking and in their approach to learning. We want to encourage them to look at the problems they missed and ask not only "What is the right answer?" but "Why didn't I get the right answer?" We help them to learn to do this with makeup exams and regrading.

The makeup is given outside of class at the end of the week following the original exam. If students are dissatisfied with their grades on the original exam, they may take it. If they do, they receive the average of the two grades on the two exams, meaning that they could lose points by choosing to take the makeup. We attempt to make the makeup exam similar in difficulty to the original but do not give the same problems.

We also explain to students that those who simply study again as they had studied for the first exam have an equal chance of going up or going down. Students who study by focusing on why they had missed problems and refining their thinking and understanding almost always improve, sometimes very substantially.[55] Typically, 25% of the class chooses to take the makeup, with 80–90% of those improving their grade significantly.

A second technique we used to focus students on thinking about their own thinking is the regrade. When we go over the exams in class, we stress that the graders had many papers to read and might not have understood a particular student's reasoning. Students who believe that they should have had a higher score on any problem may write a page with an explanation of their reasoning and argument for more points. Papers that are handed in with statements like "Please look at problem 3 again" or "I think I should have had more points on problem 2" are returned with instructions to write an explanation discussing their answer and the correct answer. The intent is to focus students on their thinking and how it compares to the solution shown in class. We hope that having to write detailed explanations can help students find where they went astray, as well as provide the instructor with an opportunity to interact one-on-one with students who have specific difficulties. Typically, about 30% of the students write requests for regrades.

Quizzes

As an attempt to give students still earlier opportunities to change their expectations about the course, we introduced weekly quizzes starting from the second week of class. These quizzes are given during the first 10 minutes of the first class of the week and typically focus on applying the processes learned in the previous week's tutorial to a new example.[37] These quizzes have the advantage of focusing students on the value of the tutorial early[56] and demonstrating that memorizing answers is not effective in this course.

The quizzes are collected and the answers given (without explanation). After class, the quizzes are graded, the answers given by each student recorded, and the quizzes handed back in the next class. That class began with a discussion of the quiz and a presentation of the distribution of answers chosen by the class. Students are asked to justify common answers and to discuss reasons for choosing them and ways for evaluating their own thinking to understand how to know they were right or wrong.

8. Coherence and synergy

We make every effort to get all the parts of the class to work together and send the same epistemological message. Tutorials and ILDs often include the same epistemological icons used in lecture. Examples in lecture are often referred to in the tutorials. And, the lecturer often speaks about issues related to in the laboratory or homework.

IV. DATA COLLECTION AND OBSERVATIONS

We collected data about student understanding and epistemologies in a variety of ways for both research and instruction including video of students working in tutorials, laboratories, and the course center as well as records and scans of student responses on clicker questions, quizzes, homework, and exams.

The project produced approximately 400 hours of videotape of students participating in tutorials, approximately 500 hours of videotape of students in laboratories, and approximately 50 hours of students working in the course center. Weekly homework, lab reports, and exams were scanned for approximately 500 students during the last four years of the project.

All students taking the project class took surveys at the beginning and at the end of the first term and at the beginning of the second term. Surveys included a mechanics conceptual survey and an attitudes/expectations survey.

We give here an overview of the evidence of progress made by students with respect to our goals. We also discuss some relevant instances in concept learning, laboratory behavior, attitudes and expectations, and student perceptions of the class.

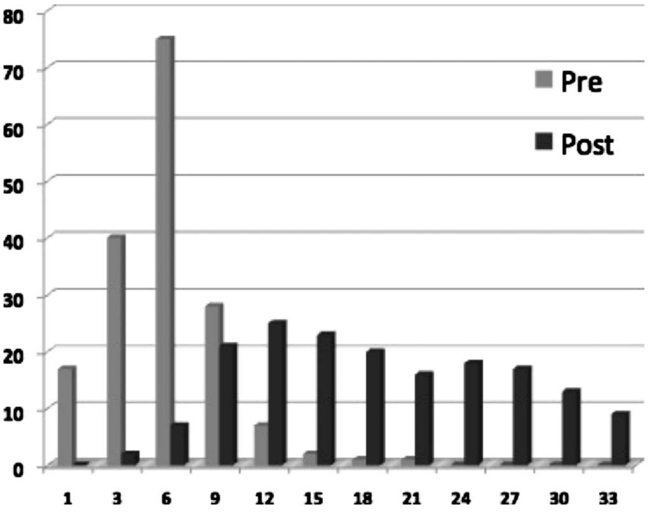


Fig. 4. Pre-post distributions of scores on the FMCE from the fall 2007 class. [Scoring using spreadsheet by M. Wittmann (Ref. 71)] The pre-post averages were 15 and 53%, and the fractional gain was $\langle g\rangle=0.44$. ($N=170$, matched.)

A. Concept learning

The primary goal of most reformed classes in the first semester of introductory physics is good conceptual learning of mechanics.[57] We placed priority on epistemological development. Although we were building our reforms on best-practices curricular materials designed and demonstrated to improve the learning of concepts, we were uncertain whether our shifted emphasis would be at the cost of the conceptual gains produced by the materials in their original form. That turned out not to be the case.

> ***It is possible to obtain strong conceptual gains in a class whose primary focus is epistemological learning.***

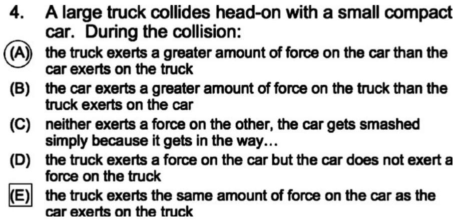
We used two widely accepted instruments: the Force Concept Inventory (FCI)[58] and the Force Motion Conceptual Evaluation (FMCE).[59] Although these tests are limited in scope and test student performance in only a single environment, they are reasonably good indicators of broader student understanding and skill development.[45]

An often used figure of merit for pre-post testing is the average fractional gain, $\langle g\rangle$, which is the fraction of the number of points the class gained compared to the fraction of number of percentage points the class could have gained.[60]

$$\langle g\rangle = \frac{(\text{post percentage average}) - (\text{pre percentage average})}{100 - (\text{pre percentage average})}. \tag{1}$$

Typical traditional high school and college classes score gains on the FCI of $\langle g\rangle \sim 0.2$, and reformed, active engagement classes score on the order of 0.35 for modest reforms and on the order of 0.6 for more extensive reforms.[43,61]

During the project, we taught the first semester course four times. A sample of our results is shown in Fig. 4. The aver-



4. A large truck collides head-on with a small compact car. During the collision:

(A) the truck exerts a greater amount of force on the car than the car exerts on the truck

(B) the car exerts a greater amount of force on the truck than the truck exerts on the car

(C) neither exerts a force on the other, the car gets smashed simply because it gets in the way...

(D) the truck exerts a force on the car but the car does not exert a force on the truck

(E) the truck exerts the same amount of force on the car as the car exerts on the truck

Fig. 5. A "split" response of a student on an FCI question.

age gains in the class ranged from 0.44 to 0.47. These results are in the range of gains shown by the stronger of the "active engagement" classes in the Hake survey[61] and are the best we have obtained at the University of Maryland.[62]

Conceptual gains are valuable, but our focus in this class is epistemological learning, to help students develop their intuition and perception that physics makes sense.

> ***It is possible to help students develop their intuitions that the physics they are learning makes sense.***

Conventional applications of conceptual surveys do not distinguish between students coming to recognize what the course considers to be the correct answers and students coming to see those answers as making sense to them. We hoped to achieve the latter and to avoid the situation of students learning to provide answers they do not personally believe.

We applied the split-survey task developed by McCaskey, Dancy, and Elby.[63] This task asks students first to "circle the answer that makes the most intuitive sense" and second to "put a square around the answer that [they] think a scientist would give." A typical example of a student response from the FCI is shown in Fig. 5. In this example, the student shows that she knows the "correct" answer, but by splitting indicates that she does not find that answer intuitive. She has not reconciled Newton's third law with her sense that the bigger (or more active) object must exert the greater force.[64]

In most semesters, only EFR delivered a reformed class, alongside two traditional sections taught at different times. Between the first and second semesters, a significant number of students transfer from one section to another, typically driven by schedule constraints.[65] In the spring semester of 2003, we gave the split task FCI to the ~200 students entering the second semester class. Approximately two-thirds of the students had taken the reformed first semester class, while the rest of students had taken a traditional one.

The results on the Newton's third law cluster of questions on the FCI (four items) for the two subsets of students are shown in Fig. 6. The students from the reformed mechanics course answered a larger percentage of the questions correctly (~85% in the reformed class compared to ~45% in the traditional), and the correct answers were not split more often (~90% in the reformed class compared to ~60% in the traditional).

These results are dramatic, with the number of "right and reconciled" answers increasing from ~25% with the traditional approach to ~75% with our reformed approach. Although there are many issues to be considered to establish a convincing research result,[66] these preliminary data suggest that something important is happening in the reformed class that is not happening in the traditional one.

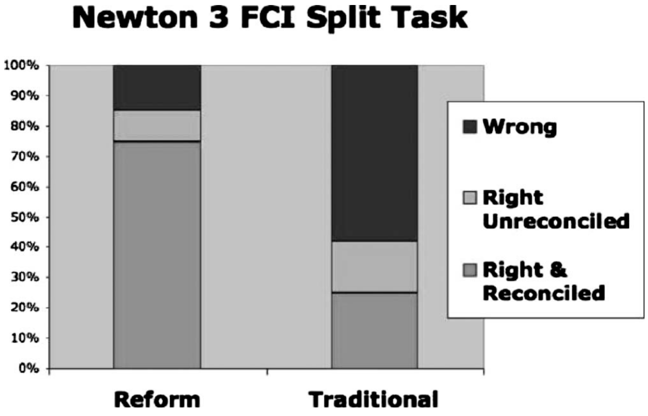


Fig. 6. Split task results on the Newton's third law cluster of the FCI. "Right and unreconciled" means the correct answer was given as the scientist's answer but a different answer was given as the intuitive one. "Right and reconciled" means the correct answer was given and not split.

B. Laboratory activities

> ***Students tended to frame our reformed laboratories as about sense-making, in contrast to simply following instructions, as students tend to frame traditional laboratories.***

To see how our students responded to our new laboratories, we analyzed videotapes of student discourse in traditional labs (taken at the beginning of the project) and in the reformed labs (taken at the end of the project).[45,47] In brief, we categorized student comments into one of four types: sense-making, logistics, off-task, and metacognitive. We found that student discourse in the reformed labs was much more likely to be comprised of sense-making—focused on the substance of the physics—than in the traditional labs: about 20% of coded comments in the reformed labs compared to 5% in the traditional labs. We also found that metacognitive statements—ones like, "I don't get this at all"—were much more likely in the reformed labs to result in some productive change in the students' behavior and reasoning. The details of these analyses are in Ref. 45.

C. Attitudes and expectations

Our results show that in the reformed classroom, students learned concepts, sensed the coherence of the physics they were learning, and spent more time in their laboratories seeking cogency. These are all measures of how students are functioning in their learning. Another way to measure students' epistemological progress is by survey. Surveys of this type include the Maryland Physics Expectations (MPEX) survey,[26] the Epistemological Beliefs Assessment for Physical Sciences (EBAPS),[67] and the Colorado Learning Attitudes About Science Survey (C-LASS).[68] The MPEX and C-LASS surveys consist of a list of statements with which the students are asked to agree or disagree on a 1–5 point scale: strongly disagree, disagree, neutral, agree, and strongly agree. The EBAPS also contains such items but adds a set of scenario items, in which students are presented with two situations and asked to decide which would be more effective in helping them learn physics. The items are clustered into several categories, including concepts, coherence, reality, mathematics, and independence. On some

Table I. Some examples of items from the attitude/expectations survey.

Item	Source	Category	Favorable polarization
When solving problems, the key is knowing the methods for addressing each particular type of question. Understanding the "big ideas" might be helpful for specially written essay questions, but not for regular physics problems.	EBAPS	Concept	Disagree
A significant problem in this course will be being able to memorize all the information I need to know.	MPEX	Coherence	Disagree
When learning a new physics topic, it's important to think about my personal experiences or ideas and relate them to the topic being analyzed.	MPEX	Coherence/Reality	Agree
Let's say a student has limited time to study, and therefore must choose between the following options. Assuming the exam will be a fair test of understanding, and assuming time pressure during the exam isn't an issue, which option should the student choose? a) Learning only a few basic formulas, but going into depth with them. b) Learning all the formulas from the relevant chapters, but not going into as much depth. c) Compromising between (a) and (b), but leaning more toward (a). d) Compromising between (a) and (b), but leaning more toward (b). e) Compromising between (a) and (b), midway between those two extremes.	EBAPS	Coherence/ Math	(a) or (c)

items, experts would prefer that the students agree with the item, on others that they would disagree. If a student's response agrees with those preferred by an expert, it is said to be favorable; otherwise, it is said to be unfavorable.

The results of these three surveys are highly consistent. Students typically enter with attitudes that agree with those preferred by experts at a level of about 50–65%. After one semester of introductory physics, these attitudes deteriorate by 5–10% or more, whether or not the class has been reformed to produce improved conceptual learning. The reality link is a particular problem. It drops by 10–20%.[26] It has been demonstrated that in a small class with a strong emphasis on epistemology, substantial gains can be obtained on such surveys.[9]

Because different populations require different surveys, we created a new survey for this study that included elements from the MPEX and the EBAPS. This survey, which we refer to as MPEX-II, is included in Ref. 37, Part 3 with the assignment of the elements to categories.[69] Four sample items, their source, their category, and their polarization (whether an "agree" response is favorable or unfavorable) are given in Table I. The main result is dramatic.

> ***It is possible to achieve significant gains on an Expectations/Attitude survey in a large lecture class without sacrificing conceptual gains.***

The pre-post matched survey results on MPEX-II ($N = 146$) are shown in Fig. 7. The class shows strong gains in the concepts and coherence categories and an insignificant gain in independence. The difficult-to-improve subcategory "reality" improved from 66% favorable to 73% favorable. The favorable percentages in each of the items in Table I improved by 30% or more. These results suggest that not only did the students improve on the functional aspects of their epistemologies in the class, but they were aware of and could recognize these changes.

D. Student perceptions of the class

In addition to the pre-post surveys, we have additional information about the student perceptions of our reforms: a survey carried out by the university administration that was done during one of our reform classes and some particular instances that illustrate phenomena we observed with a larger number of students.

1. A serendipitous external evaluation: The CORE survey

The University of Maryland has a series of distribution requirements known as CORE.[70] Courses approved for CORE are intended to help students gain "a strong and broadly based education," to "introduce the great ideas and controversies in human thought and experience," and to provide "a strong foundation for…life-long learning."[70] Algebra-based physics has been approved for decades as suitable for meeting the science component of CORE. Every few years, the university's central administration carries out an evaluation of the CORE classes, one of which took place during this project.

The evaluation consisted of a survey with eight items answered on a scale of 1 to 5 ("not at all" to "a great deal").

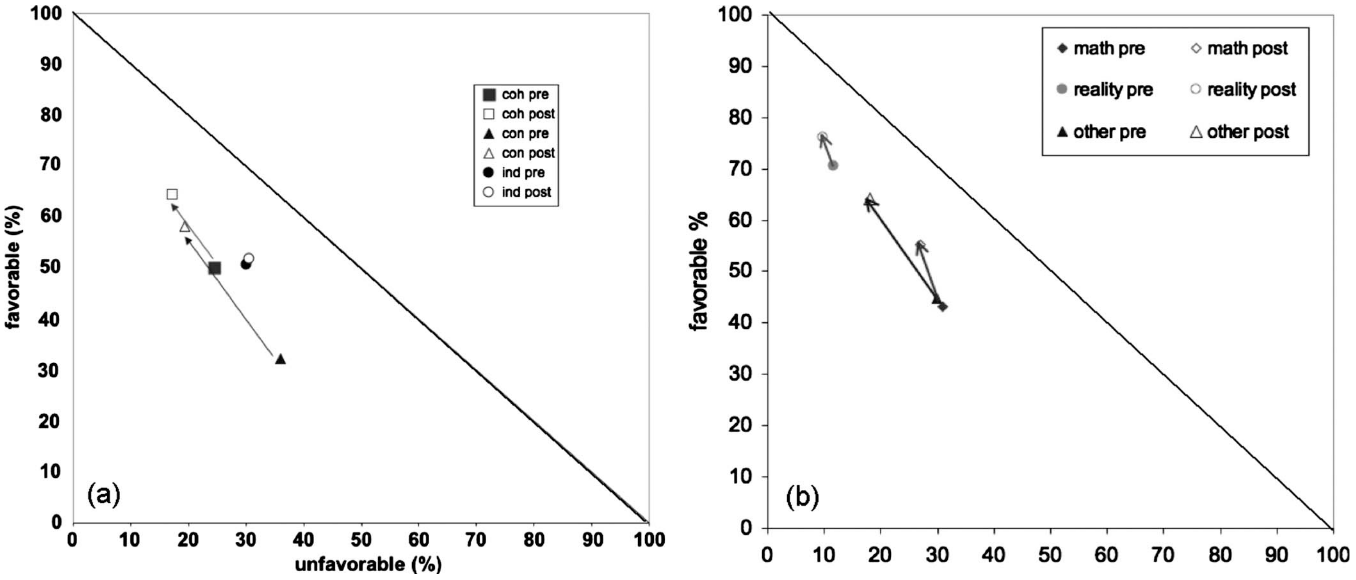


Fig. 7. The results on the MPEX-II in the first semester of one of our reformed classes. Other semesters (and our other instructor) were similar. (a) The main categories of concepts, coherence, and independence. (b) Subcategories of concepts: math, reality, and other.

The items were of the form: “To what extent has this course been intellectually stimulating?” and “To what extent have the writing assignments and/or examinations in this course given you opportunities to think carefully and critically?” (The full survey is given in Ref. 37, Part 4.) The survey was also given to another lecture section at the same time that was taught using the traditional method. The traditional section scored sufficiently strongly to retain CORE approval (results between 2.5 and 4.0 on the various items). These results agreed to within 0.2 item-by-item with the survey done in the class five years earlier with a different traditional instructor. The results in our reformed class were almost a full point higher than in the traditional section for each item. These results are encouraging, especially given that the survey was developed independently and we were unaware that the survey was to be done and therefore did not intentionally orient the class toward the issues being measured.

2. Resistance: It can take a while for students to get used to the approach

Many of the students in this class were juniors and seniors, and about one-third were pre-meds. Many had considerable experience in science classes, and many had outstanding academic records. They thought they knew what was demanded of them, and they thought they knew how to cope with those demands. When we told them at the beginning of the term that we were “changing the rules,” some got nervous, and some got angry.

In one case, one of us (EFR) heard a report from a TA that one student, “Karen” (pseudonym), was being especially resistant. A few days later, EFR was called into the Chair’s office to respond to a parent’s complaint. After reassuring the Chair that the class was under control, he passed a message to Karen through the TA to come and see him. When she arrived, her body language suggested a mix of emotions—nervousness, anger, and uncertainty. The instructor listened to the student’s complaint calmly and was reassuring. She was a 4.0 pre-med who felt confident of being able to succeed in a traditional class and was unsure of being able to meet our new expectations. He reassured her that we understood, that we had prepared instructional resources to help her, and that his door was always open.

Karen achieved an A on the first exam and continued to perform strongly, earning an A both semesters. At the end of the year, EFR received a letter from her with the following comment: “Your class was one of the most interesting and beneficial classes I’ve taken at the University—I improved my thinking skills, creativity, and teamwork skills, not to mention that I learned a lot of Physics! Your style of teaching was one that I feel lucky to have benefited from…and my mom thought so too!…So I’d like to first express my gratitude to you for providing that experience.”

3. Some individuals undergo significant changes in their approach to science

One extreme case in our other instructor’s class (DH) was a student, “Louis,” who failed the first midterm exam with a score of 36%. This failure prompted him to meet with the professor and ask what he was doing wrong. The instructor advised the student to try to make sense of the material by considering “how he would explain it to a ten-year-old.” Louis’s score on the make-up exam jumped to 84%, the highest on the make-up and near the top of the original distribution. In a videotaped interview with a researcher not involved in teaching the course, he explained that his interaction with the professor prompted him to change his approach to studying, from “memorizing the book [and] every word of those homework solutions” to trying to “write down an explanation like to a ten-year-old” using analogies to everyday ideas.

Six months after the course was over, Louis wrote to tell the professor, “since I’ve taken your class, I have a 4.0 GPA, compared to a much lower GPA before your class. I think this increase in GPA has a lot to do with the things I learned in your class—not about physics, but about learning in general.” Louis explained why the advice to “explain it to a ten-year-old” had been effective for him: He had experience tutoring, both children and peers, and at the time he was enrolled in the course he was working as a tutor, using strategies of trying to connect to what his tutees already knew. The professor's advice made him more aware of what he was doing as a learner and connected to epistemological resources that he possessed but was not previously making use of in his physics class.[7]

V. DISCUSSION AND CONCLUSIONS

We have summarized a set of reforms for algebra-based physics intended to explicate the epistemological elements of the implicit curriculum and to provide students explicit instruction in learning how to think about science and understand the process of scientific reasoning. The transformations of the instructional environment were built on existing best-practices curricula that had been demonstrated to provide strong conceptual learning but to have little effect on expectations and attitudes. We reformed these elements based on a theory of thinking and learning, the resources framework. This framework focuses attention on the resources, both conceptual and epistemological, that students bring to class.

Students responded well to these curricular transformations, demonstrating both strong conceptual learning and increased ability to use and articulate the need for concepts, coherence, and cogency. As with any major reform, especially one where student expectations are not met, it is not sufficient to simply introduce the reforms by using the transformed materials. Considerable effort is needed to help students make sense of and become comfortable with them. Our experience suggests three important ways to do so: attend to student framing, be consistent, and restrict the content appropriately.

A. Attend to student framing

Students frame the way they think about our class based on their previous experience in other science classes. If these courses involved an inappropriate implicit epistemological curriculum (for example, rewarded memorization and discouraged sense-making[25]), students might not bother to pay attention to statements that this course is going to be different. Despite explicit statements on the first day of class, detailed handouts explaining the goals of the course, and repeated statements in class, many of our students are primed to ignore anything we say that they do not interpret as direct content.[72] Despite a statement in the main handout that homework was the most important learning activity and repeated mentions of it in the class, one student commented on the anonymous faculty evaluation at the end of term, "You won't believe this, but I actually learned the most in this class doing homework!" Despite explicit statements in both the main handout and in class that homework is not graded for feedback and you have to read the posted solutions, some students reported that "TA's didn't grade for correct answers on the HW's so you went through the semester THINKING you had the right idea, till you got your exam back and you were wrong."

For an epistemologically oriented class it appears necessary to be continually working to help the students reframe their understanding of what success in the class entails—articulating the class expectations and epistemological goals and listening to (and demanding) their feedback on what is happening.

B. Be consistent

Traditional courses send unspoken messages about the implicit epistemological curriculum; our reformed course sends very different and (hopefully) explicit ones. One of us (EFR) had been teaching large lecture classes in a traditional mode for nearly twenty-five years and often found himself slipping "off message" unintentionally. When one of us (DH) taught a semester of the class with traditional rather than reformed laboratories, he found smaller MPEX-II gains than when he taught with a reformed lab.

C. Restrict content appropriately

One challenge that faculty considering a reformed class often make is "What do you have to leave out?" The idea that one has to cover a particular set of material, whether or not the students understand it, seems peculiar, but it is widespread. An approach that is more appropriate to our goals is to "uncover a little rather than cover a lot."[73] Having large blocks of material in a class that the students cannot expect to understand means lots of material that they will have to memorize. That might not be a problem, but because students have intuitive epistemologies, the effect may not be restricted to that particular material. Students may learn or have reinforced the idea that physics does not make sense and apply that idea to material they could otherwise have understood. We attempted to remove topics that appeared to lead even our top students into confusion despite our best attempts with whatever best-practice materials we could bring to bear. We determined which topics fit this description by having final exam questions that included enough open ended and explanatory components to give us some insight into student thinking on these topics. In this way we were lead to eliminate topics such as heat engines, magnetic induction (Faraday's law), Gauss's law, the details of electromagnetic waves, and much of modern physics.[74] Some of this elimination we did with great regret, as we believe it is important for the biology majors to know about these topics. However, we accepted the idea that we could not teach everything and that it was epistemologically more effective to have students learn topics in physics that they could genuinely understand rather than be exposed to topics that were interesting but that they could not master in the limited time allotted.

These decisions are highly dependent on instructional tools and environments. It is possible that someone will learn how to teach Faraday's law in a way that this population can make sense of with only a few hours of instruction. When this occurs, it will be appropriate to rethink our selection of content.

VI. CONCLUSION

In any class, we teach an implicit epistemological curriculum. What many students learn about these issues of the nature of scientific knowledge and what it means to learn and understand science might be the most important knowledge they take away from our class.[75] When these lessons are tacit, inexplicit, and unevaluated, students may learn the opposite of what we intend. Our reforms in algebra-based physics are an illustration of how an epistemological curriculum can be analyzed, explicated, and evaluated for a particular population of students. A critical issue is an understanding of how students' intuitive epistemologies play a role in their learning. Attending to that issue may prove to be of value in other classes as well.

ACKNOWLEDGMENTS

We are grateful to the members of the University of Maryland Physics Education Research Group who were responsible for carrying out the research and development reported here, in particular, faculty members Andy Elby, Laura Lising, and Rachel Scherr and graduate students Paul Gresser, Ray Hodges, Rebecca Lippmann, Tim McCaskey, Rosemary Russ, and Jonathan Tuminaro. This material is based on work supported by the U.S. National Science Foundation under Awards No. REC-00-87519 and REC-04-40113. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation. One of us (EFR) would like to thank the University of Colorado and JILA for sabbatical support while this manuscript was being written.

[1] NSF Grant REC-008-7519, *Learning How to Learn Science: Physics for Bioscience Majors*.

[2] B. Alberts, "The cell as a collection of protein machines: Preparing the next generation of molecular biologists," Cell **92**, 291–294 (1998).

[3] H. Varmus, "The impact of physics on biology and medicine," APS News 8 (August/September, 1999).

[4] *Bio 2010: Transforming Undergraduate Education for Future Research Biologists* (National Academy of Sciences Press, Washington, DC, 2003).

[5] David Hammer, "Student resources for learning introductory physics," Am. J. Phys. **68**(7), S52–S59 (2000).

[6] E. F. Redish, "A Theoretical Framework for Physics Education Research: Modeling Student Thinking," in *Proceedings of the International School of Physics, "Enrico Fermi" Course CLVI*, edited by E. F. Redish and M. Vicentini (IOS Press, Amsterdam, 2004), pp. 1–63.

[7] D. Hammer, A. Elby, R. E. Scherr, and E. F. Redish, "Resources, Framing, And Transfer," in *Transfer of Learning from a Modern Multidisciplinary Perspective*, edited by Jose Mestre (Information Age Publishing, Charlotte, NC, 2004), Chap. 3.

[8] D. Hammer, "Discovery learning and discovery teaching," Cogn. Instruct. **15**, 485–529 (1997).

[9] A. Elby, "Helping physics students learn how to learn," Am. J. Phys. **69**(7), S54–S64 (2001).

[10] D. Hammer and A. Elby, "Tapping students' epistemological resources," J. Learn. Sci. **12**(1), 53–91 (2003).

[11] The notion of a "hidden curriculum" has sometimes been used in educational thought to describe problematic lessons students can take away from the structures and practices of schools, with an emphasis on social values. See, for example, B. R. Snyder, *The Hidden Curriculum* (Knopf, 1971), instructor ed. and H. A. Giroux, *Teachers as Intellectuals: Toward a Critical Pedagogy of Learning* (Bergin & Garvey, New York, 1988) and *Theory and Resistance in Education: Towards a Pedagogy for the Opposition* (Bergin & Garvey, New York, 2001). Here we focus specifically on student epistemologies, which are undoubtedly related to social values, but we do not pursue the relationship.

[12] E. F. Redish, *Teaching Physics with the Physics Suite* (Wiley, New York, 2003), Chap. 2.

[13] In earlier work we have referred to some of these issues as issues of student *expectations*. We now view these expectations as reflections of students' context-dependent epistemological understandings of the nature of the knowledge they are learning in a particular situation and what they think they have to do to learn it.

[14] L. Lising and A. Elby, "The impact of epistemology on learning: A case study from introductory physics," Am. J. Phys. **73**, 372–382 (2005).

[15] E. F. Redish, R. E. Scherr, and J. Tuminaro, "Reverse engineering the solution of a 'simple' physics problem: Why learning physics is harder than it looks," Phys. Teach. **44**(5), 293–300 (2006).

[16] J. Tuminaro and E. F. Redish, "Elements of a cognitive model of physics problem solving: Epistemic games," Phys. Rev. ST Phys. Educ. Res. **3**, 020101-1–22 (2007).

[17] R. E. Scherr, "Gesture analysis for physics education researchers," Phys. Rev. ST Phys. Educ. Res. **4**(1), 10101-1–9 (2008).

[18] R. E. Scherr and D. Hammer, "Student behavior and epistemological framing: Examples from collaborative active-learning activities in physics," Cogn. Instruct. (in press).

[19] G. Hatano and K. Inagaki, "Two Courses of Expertise," in *Child Development and Education in Japan*, edited by H. Stevenson, H. Azuma, and K. Hakuta (Freeman, New York, 1986), pp. 262–272.

[20] A. Einstein, "Physics and Reality," J. Franklin Institute **221**(3), 349–382 (1936).

[21] A. A. diSessa, "Toward an epistemology of physics," Cogn. Instruct. **10**, 105–225 (1993).

[22] J. Clement, D. Brown, and A. Zeitsman, "Not all preconceptions are misconceptions: Finding 'anchoring conceptions' for grounding instruction on students' intuitions," Int. J. Sci. Educ. **11**, 554–565 (1989).

[23] Michael McCloskey, "Naïve Theories of Motion," in *Mental Models*, edited by Dedre Gentner and Albert L. Stevens (Erlbaum, Hillsdale, NJ, 1983), pp. 299–324.

[24] K. A. Strike and G. J. Posner, "A Revisionist Theory of Conceptual Change," in *Philosophy of Science, Cognitive Psychology, and Education Theory and Practice*, edited by R. A. Duschl and R. J. Hamilton (SUNY Press, New York, 1992), pp 147–176.

[25] D. Hammer, "Two approaches to learning physics," Phys. Teach. **27**, 664–671 (1989); D. Hammer, "Epistemological beliefs in introductory physics," Cogn. Instruct. **12**(2), 151–183 (1994).

[26] E. F. Redish, J. M. Saul, and R. N. Steinberg, "Student expectations in introductory physics," Am. J. Phys. **66**, 212–224 (1998).

[27] As a result of our project, some of the sections are currently more or less nontraditional. For example, about half of the sections now use some student response devices in lecture and nontraditional laboratories and tutorials. As of this writing, about half of the sections remain as described here.

[28] From the University of Maryland online course catalog, spring term 2007 (www.sis.umd.edu/bin/soc?term=200701e&crs=PHYS).

[29] The course was taught most often by the first author, whose version of the course is the principal basis of this article.

[30] J. Nussbaum and S. Novick, "Alternative frameworks, conceptual conflict and accommodation: Toward a principled teaching strategy," Instr. Sci. **11**(2), 183–200 (1983).

[31] L. C. McDermott, "Millikan Lecture 1990: What we teach and what is learned—Closing the gap," Am. J. Phys. **59**, 301–315 (1991).

[32] E. Mazur, *Peer Instruction: A User's Manual* (Prentice Hall, New York, 1997); C. Crouch, J. Watkins, A. Fagan, and E. Mazur, "Peer Instruction: Engaging students one-on-one all-at-once," in *Reviews of Research-Based Reform Curricula in Introductory Physics*, edited by E. F. Redish and P. Cooney (www.compadre.org/PER/items/detail.cfm?ID=4990).

[33] D. Sokoloff and R. Thornton, *Interactive Lecture Demonstrations in Introductory Physics* (Wiley, New York, 2004); D. R. Sokoloff and R. K. Thornton, "Using interactive lecture demonstrations to create an active learning environment," Phys. Teach. **35**, 340–347 (1997).

[34] These icons are available in Ref. 37 (part 1).

[35] D. Hammer, "Two approaches to learning physics," Phys. Teach. **27**, 664–671 (1989).

[36] D. Schachter, *Seven Sins of Memory: How the Mind Forgets and Remembers* (Houghton-Mifflin, Boston, MA, 2001).

[37] See EPAPS Document No. E-AJPIAS-77-007906 for supplementary online materials. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.

[38] Available online at (www.physics.umd.edu/perg/ILD.htm).

[39] F. M. Goldberg and L. C. McDermott, "An investigation of student understanding of the real image formed by a converging lens or concave mirror," Am. J. Phys. **55**, 108–119 (1987).

[40] R. M. Goertzen, R. E. Scherr, and A. Elby, "Indictors of Understanding: What TAs Listen For in Student Responses," *2008 Physics Education Research Conference, AIP Conf. Proc.*, 1064, 119–122 (2008).

[41] L. McDermott, P. Shaffer, and the University of Washington Physics Education Group, *Tutorials in Introductory Physics* (Prentice Hall, New York, 2002).

[42] M. Wittmann, R. Steinberg, and E. Redish, *Activity-Based Tutorials* (Wiley, New York, 2004).

[43] E. F. Redish, J. M. Saul, and R. N. Steinberg, "On the effectiveness of active-engagement microcomputer-based laboratories," Am. J. Phys. **65**, 45–54 (1997).

[44] The creation of such a pair requires a good knowledge of both the research database (to identify common misconceptions) and students productive resources (to generate contexts that will facilitate students in spontaneously proposing correct answers).

[45] R. Lippmann, "Students' understanding of measurement and uncertainty in the physics laboratory: Social construction, underlying concepts, and quantitative analysis," PhD dissertation, University of Maryland, College Park, MD (2003), (www.physics.umd.edu/perg/dissertations/Lippmann).

[46] F. Reif and M. St. John, "Teaching physicists' thinking skills in the laboratory," Am. J. Phys. **47**, 950–957 (1979).

[47] R. L. Kung and C. Linder, "Metacognitive activity in the physics student laboratory: is increased metacognition necessarily better?" Metacognition and Learning **2**(1), 41–56 (2008).

[48] R. Lippmann Kung, "Teaching the concepts of measurement: An example of a concept-based laboratory course," Am. J. Phys. **73**(8), 771–777 (2005).

[49] P. Gresser, "A Study Of Social Interaction And Teamwork In Reformed Physics Laboratories," PhD dissertation, University of Maryland, College Park, MD (2006), (www.physics.umd.edu/perg/dissertations/Gresser).

[50] A. H. Schoenfeld, *Mathematical Problem Solving* (Academic Press, New York, 1985).

[51] Reference 12, Chap. 4.

[52] The "Thinking Problems in Physics" collection is available at (www.physics.umd.edu/perg/abp/TPProbs/ProbSubjs.htm). Many of these problems have solutions that are password protected. The password is available to instructors upon request to redish@umd.edu.

[53] Unfortunately, we have no data on how many students acted on this admonition.

[54] For details see Ref. 16.

[55] We only have informal experience to this effect and have not conducted a systematic study.

[56] The traditional tutorial model (Ref. 41) relies on an ungraded cognitive-conflict pretest and on a tutorial question in each mid-semester exam to achieve similar goals. Our post-test graded model was a better fit to our epistemological goals.

[57] See, for example, Ref. 12, Chaps. 7–9.

[58] D. Hestenes, M. Wells, and G. Swackhamer, "Force concept inventory," Phys. Teach. **30**, 141–158 (1992).

[59] R. K. Thornton and D. R. Sokoloff, "Assessing student learning of Newton's laws: The force and motion conceptual evaluation," Am. J. Phys. **66**(4), 228–351 (1998).

[60] This measure is typically used to permit the comparison of classes with different pre-test scores. Some care must be taken, as the score is particular to the specific test used and may distort the result when either the pre—or post-test averages are high, producing end effects. We are grateful to Robert Mislevy (private communication) for this comment.

[61] R. Hake, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," Am. J. Phys. **66**, 64–74 (1998).

[62] Although the other classes we have tested at Maryland were all calculus-based, the Hake survey shows that for the FCI in the mid range (pre-test scores from 20 to 60%), similarly structured courses tend to produce similar fractional gains in high school, algebra-based, and calculus-based physics classes. We have compared to our own results rather than to the full Hake set, because the Hake set contained self-reported results from many classes and may have mixed together classes with distinct styles of reform. For our own classes, we know exactly what was done in each case.

[63] T. McCaskey, M. Dancy, and A. Elby, "Effects on Assessment Caused by Splits Between Belief and Understanding," in *2003 Physics Education Research Conference*, edited by J. Marx, K. Cummings, and S. Franklin, AIP Conf. Proc. **720**, 37–40 (2004).

[64] L. Bao, K. Hogg, and D. Zollmann, "Model analysis of fine structures of student models: An example with Newton's third law," Am. J. Phys. **70**, 766–778 (2002).

[65] We cannot rule out the possibility that students left or entered the reform class because of personal preferences associated with the style of teaching.

[66] We did not conduct a randomized controlled comparison of our course with others for several reasons. One was logistical: Student course selections are driven largely by schedule constraints, and to have the same population would require a second, independent "control" lecture meet at the same time as the "treatment" sections, which would not be possible under the current system. A second reason is ethical: To conduct a randomized trial would require that we require some students who would prefer otherwise to enroll in a traditional section, when there is extensive evidence that traditional approaches are problematic. Thus, this comparison is clouded by selection effects, both that the schedule constraints affect the population of students in each section, as well as the effects of students' choices who are not so constrained (in both directions). Some of these issues are addressed in T. McCaskey, PhD Thesis, U. of Maryland, forthcoming.

[67] (www2.physics.umd.edu/~elby/EBAPS/home.htm).

[68] W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman, "New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey," Phys. Rev. ST Phys. Educ. Res. **2**, 010101-1–14 (2006).

[69] A paper on the construction, validation, and detailed results of the survey is in preparation, A. Elby and T. McCaskey (private communication).

[70] From the Undergraduate Catalog (www.umd.edu/catalog/0607/Chapter5.pdf).

[71] M. Wittmann, (perlnet.umephy.maine.edu/materials/template_fmce.zip).

[72] Such selective attention is well documented in cognitive psychology. See, for example, D. J. Simons and C. F. Chabris, "Gorillas in our midst: Sustained inattentional blindness for dynamic events," Perception **28**, 1059–1074 (1999).

[73] Attributed to Victor Weisskopf; Union of Concerned Scientists, "Victor Weisskopf remembered," (http://go.ucsusa.org/ucs/members/page.cfm?pageID=1015).

[74] We did not attempt to include special relativity, believing that, although it is exciting and interesting, few of the students would be able to get more than a "gee whiz" insight in the limited time we had available. The reasons many faculty give for including relativity is that it shows that physics has surprises for us if we pay careful attention to experiment and inference. Although these are valuable epistemological messages that fit well with our course, the same messages can be taught in Newtonian mechanics and ray optics much more successfully in the time that is available and in a way that can excite students almost as much. But in the more classical cases, students also take the message that these surprises are something they can actually understand.

[75] "Education is what survives when what has been learned has been forgotten." B. F. Skinner, quoted in New Scientist (21 May 1964).