WHY NOT TRY

A SCIENTIFIC APPROACH TO SCIENCE EDUCATION?

By Carl Wieman

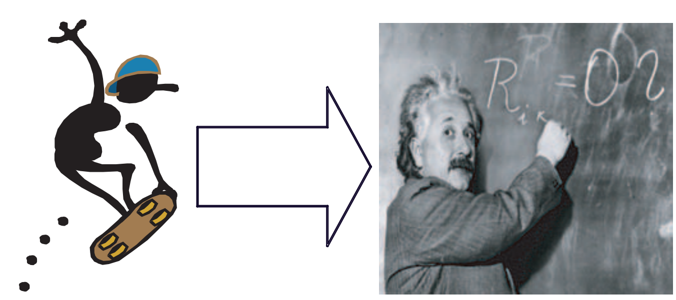
The purpose of science education is no longer simply to train that tiny fraction of the population who will become the next generation of scientists. We need a more scientifically literate populace to address the global challenges that humanity now faces and that only science can explain and possibly mitigate, such as global warming, as well as to make wise decisions, informed by scientific understanding, about issues such as genetic modification. Moreover, the modern economy is largely based on science and technology, and for that economy to thrive and for individuals within it to be successful, we need technically literate citizens with complex problem-solving skills.

Carl Wieman, recipient of the Nobel Prize in physics in 2001 and the Carnegie Foundation's U.S. University Professor of the Year Award in 2004, currently directs the Carl Wieman Science Education Initiative at the University of British Columbia and the Colorado Science Education Initiative. Prior to joining UBC, he served on the faculty at the University of Colorado from 1984 to 2006 as a distinguished professor of physics and presidential teaching scholar. This article is adapted from his lecture at the Carnegie Foundation's Centennial celebration at the Library of Congress.

In short, we now need to make science education effective and relevant for a large and necessarily more diverse fraction of the population.

What do I mean by an effective education in science? I believe a successful science education transforms how students think, so that they can understand and use science like scientists do. (See Figure 1). But is this kind of transformation really possible for a large fraction of the total population?

Figure 1. Transporting student thinking from novice to expert.



The hypothesis that I and others have advanced is that it *is* possible, but only if we approach the teaching of science like a science. That means applying to science teaching the practices that are essential components of scientific research and that explain why science has progressed at such a remarkable pace in the modern world.

The most important of these components are:

• Practices and conclusions based on objective data rather than—as is frequently the case in education—anecdote or tradition. This includes using the results of prior research, such as work on how people learn.

• Disseminating results in a scholarly manner and copying and building upon what works. Too often in education, particularly at the postsecondary level, everything is reinvented, often in a highly flawed form, every time a different instructor teaches a course. (I call this problem "reinventing the square wheel.")

• Fully utilizing modern technology. Just as we are always looking for ways to use technology to advance scientific research, we need to do the same in education.

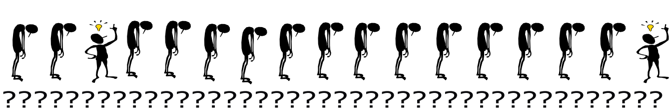
These three essential components of all experimental scientific research (and, not incidentally, of the scholarship of teaching and learning) can be equally valuable in science education. Applied to the teaching of science, they have the capability to dramatically improve both the effectiveness and the efficiency of our educational system.

The Learning Puzzle

When I first taught physics as a young assistant professor, I used the approach that is all too common when someone is called upon to teach something. First I thought very hard about the topic and got it clear in my own mind. Then I explained it to my students so that they would understand it with the same clarity I had.

At least that was the theory. But I am a devout believer in the experimental method, so I always measure results. (See Figure 2.) And whenever I made any serious attempt to determine what my students were learning, it was clear that this approach just didn't work. An occasional student here and there might have understood my beautifully clear and clever explanations, but the vast majority of students weren't getting them at all.

Figure 2. Student reaction to my brilliantly clear explanations.



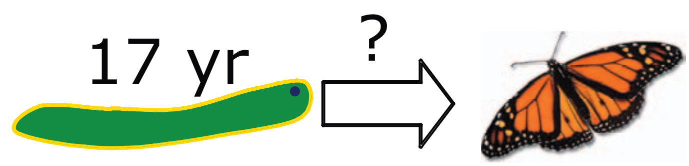
For many years, this failure of students to learn from my explanations remained a frustrating puzzle to me, as I think it is for many diligent faculty members. What eventually led me to understand it was that I was encountering the even bigger puzzle of my graduate students.

I have conducted an extensive research program in atomic physics over many years that has involved many graduate students, on whose professional development I have spent a lot of time and thought. And over the years I became aware of a consistent pattern. New graduate students would come to work in my laboratory after 17 years of extraordinary success in classes, but when they were given research projects to work on, they were clueless about how to proceed. Or worse—often it seemed that they didn't even really understand what physics was.

But then an amazing thing happened: After just a few years of working in my research lab, interacting with me and the other students, they were transformed. I'd suddenly realize they were now expert physicists, genuine colleagues. If this had happened only once or twice it would have just seemed an oddity, but I realized it was a consistent pattern. So I decided to figure it out.

One hypothesis that occurred to me, as it has to many other research advisors who have observed similar transformations, is that the human brain has to go through a 17-year "caterpillar" stage before it is suddenly transformed into a physicist "butterfly." (See Figure 3.) But I wasn't satisfied with that explanation, so I tackled it like a science problem. I started studying the research on how people learn, particularly how they learn science, to see if it could provide a more satisfactory explanation of the pattern. Sure enough, the research did have another explanation to offer that also solved the earlier puzzle of why my classroom teaching was ineffective.

Figure 3. Brain-development possibility: 17 years as intellectual caterpillar before transformation into physicist butterfly?



Research on Learning

In a traditional science class, the teacher stands at the front of the class lecturing to a largely passive group of students. Those students then go off and do back-of-the-chapter homework problems from the textbook and take exams that are similar to those exercises.

The research has several things to say about this pedagogical strategy, but I'll focus on three findings—the first about the retention of information from lecture, the second about understanding basic concepts, and the third about general beliefs regarding science and scientific problem-solving. The data I discuss were mostly gathered in introductory college physics courses, but these results are consistent with those of similar studies done in other scientific disciplines and at other grade levels. This is understandable, because they are consistent with what we know about cognition.

Retaining Information

Lectures were created as a means of transferring information from one person to many, so an obvious topic for research is the retention of the information by the many. The results of three studies—which can be replicated by any faculty member with a strong enough stomach—are instructive. The first is by Joe Redish, a highly regarded physics professor at the University of Maryland. Even though the students thought his lectures were wonderful, Joe wondered how much they were actually learning. So he hired a graduate student to grab students at random as they filed out of class at the end of the lecture and ask, "What was the lecture you just heard about?" It turned out that the students could respond with only with the vaguest of generalities.

Zdeslav Hrepic, N. Sanjay Rebello, and Dean Zollman at Kansas State University carried out a much more structured study. They asked 18 students from an introductory physics class to attempt to answer six questions on the physics of sound and then, primed by that experience, to get the answers to those questions by listening to a 14-minute, highly polished commercial videotaped lecture given by someone who is supposed to be the world's most accomplished physics lecturer. On most of the six questions, no more than one student was able to answer correctly.

In a final example, a number of times Kathy Perkins and I have presented some non-obvious fact in a lecture along with an illustration, and then quizzed the students 15 minutes later on the fact. About 10 percent usually remember it by then. To see whether we simply had mentally deficient students, I once repeated this experiment when I was giving a departmental colloquium at one of the leading physics departments in the United States. The audience was made up of physics faculty members and graduate students, but the result was about the same—around 10 percent.

Given that there are thousands of traditional science lectures being given every day, these results are quite disturbing. Do these findings make sense? Could this meager transfer of information in lectures be a generic problem?

These results do indeed make a lot of sense and probably are generic, based on one of the most well-established—yet widely ignored—results of cognitive science: the extremely limited capacity of the short-term working memory. The research tells us that the human brain can hold a maximum of about seven different items in its short-term working memory and can process no more than about four ideas at once. Exactly what an "item" means when translated from the cognitive science lab into the classroom is a bit fuzzy. But the number of new items that students are expected to remember and process in the typical hour-long science lecture is vastly greater. So we should not be surprised to find that students are able to take away only a small fraction of what is presented to them in that format.

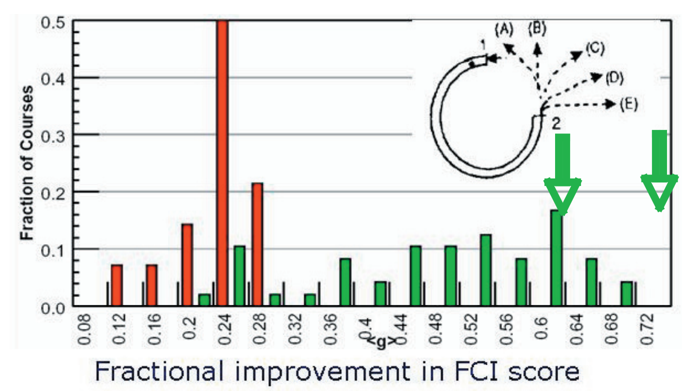
Understanding Basic Concepts

We physicists believe that one of the great strengths of physics is that it has a few fundamental concepts that can be applied very widely. This has inspired physics-education researchers to study how well students are actually learning the basic concepts in their physics courses, particularly at the introductory level.

These researchers have created some good assessment tools for measuring conceptual understanding. Probably the oldest and most widely used of these is the Force Concepts Inventory (FCI) (see Hestenes, 1992 in "Resources" below). This instrument tests students' mastery of the basic concepts of force and motion, which are covered in every first-semester postsecondary physics course. The FCI is composed of carefully developed and tested questions that usually require students to apply the concepts of force and motion in a real-world context, such as explaining what happens when a car runs into a truck. The FCI—now administered in hundreds of courses annually—normally is given at the beginning and end of the semester to see how much students have learned during the course.

Richard Hake compiled the FCI results from 14 different traditional courses and found that in the traditional lecture course, students master no more than 30 percent of the key concepts that they didn't already know at the start of the course (See Figure 4). Similar sub-30-percent gains are seen in many other unpublished studies and are largely independent of lecturer quality, class size, and institution. The consistency of those results clearly demonstrates that the problem is in the basic pedagogical approach: The traditional lecture is simply not successful in helping most students achieve mastery of fundamental concepts. Pedagogical approaches involving more interactive engagement of students show consistently higher gains on the FCI and similar tests.

Figure 4.



Source: Plot from R. Hake, "A six-thousand-student survey," AJP *66, 64-74 (1998).*

Affecting Beliefs

Students believe certain things about what physics is and how one goes about learning the discipline, as well as how one solves problems in physics. If you interview a lot of people, you find that their beliefs lie on a spectrum that ranges from "novice" to "expert." My research group and others have developed survey instruments that can measure where on this scale a person's beliefs lie.

What do we mean by a "novice" in this context? Adapting the characterization developed by David Hammer, novices see the content of physics instruction as isolated pieces of information—handed down by an authority and disconnected from the world around them—that they can only learn by memorization. To the novice, scientific problem-solving is just matching the pattern of the problem to certain memorized recipes.

Experts—i.e., physicists—see physics as a coherent structure of concepts that describe nature and that have been established by experiment. Expert problem-solving involves employing systematic, concept-based, and widely applicable strategies. Since this includes being applicable in completely new situations, this strategy is much more useful than the novice problem-solving approach.

Once you develop the tools to measure where people's beliefs lie on this expert-to-novice scale, you can see how students' beliefs change as a result of their courses. What you would expect, or at least hope, is that students would begin their college physics course somewhere on the novice side of the scale and that after completing the course they would have become more expert-like in their beliefs.

What the data say is just the opposite. On average, students have more novice-like beliefs after they have completed an introductory physics course than they had when they started; this was found for nearly every introductory course measured. More recently, my group started looking at beliefs about chemistry. If anything, the effect of taking an introductory college chemistry course is even worse than for taking physics.

So we are faced with another puzzle about traditional science instruction. This instruction is explicitly built around teaching concepts and is being provided by instructors who, at least at the college level, are unquestionably experts in the subject. And yet their students are not learning concepts, and they are acquiring novice beliefs about the subject. How can this be?

Research on learning once again provides answers. Cognitive scientists have spent a lot of time studying what constitutes expert competence in any discipline, and they have found a few basic components. The first is that experts have lots of factual knowledge about their subject, which is hardly a surprise. But in addition, experts have a mental organizational structure that facilitates the retrieval and effective application of their knowledge. Third, experts have an ability to monitor their own thinking ("metacognition"), at least in their discipline of expertise. They are able to ask themselves, "Do I understand this? How can I check my understanding?"

A traditional science instructor concentrates on teaching factual knowledge, with the implicit assumption that expert-like ways of thinking about the subject come along for free or are already present. But that is not what cognitive science tells us. It tells us instead that students need to develop these different ways of thinking by means of extended, focused mental effort. Also, new ways of thinking are always built on the prior thinking of the individual, so if the educational process is to be successful, it is essential to take that prior thinking into account.

This is basic biology. Everything that constitutes "understanding" science and "thinking scientifically" resides in the long-term memory, which is developed via the construction and assembly of component proteins. So a person who does not go through this extended mental construction process simply cannot achieve mastery of a subject.

When you understand what makes up expert competence and how it is developed, you can see how cognitive science accounts for the classroom results that I presented earlier. Students are not learning the scientific concepts that enable experts to organize and apply the information of the discipline, nor are they being helped to develop either the mental organizational structure that facilitates the retrieval and application of that knowledge or a capacity for metacognition. So it makes perfect sense that they are not learning to think like experts, even though they are passing science courses by memorizing facts and problem-solving recipes.

To the novice, scientific problem-solving is just matching the pattern of the problem to certain memorized recipes.

Improved Teaching and Learning

If we now return to the puzzle of my graduate students—why their first 17 years of education seemed so ineffective, while a few years of doing research turned graduate students into expert physicists—we see that the first part of the mystery is solved: Those traditional science courses did little to develop expert-like thinking about physics. But why is working in a research lab so different?

A lot of educational and cognitive research can be reduced to this basic principle: People learn by creating their own understanding. But that does not mean they must or even can do it without assistance. Effective teaching facilitates that creation by getting students engaged in thinking deeply about the subject at an appropriate level and then monitoring that thinking and guiding it to be more expert-like.

When you put it in those terms, you realize that this is exactly what all my graduate students are doing 18 or 20 hours a day, seven days a week. (Or at least that is what they claim—the reality is a bit less.) They are focused intently on solving real physics problems, and I regularly probe how they're thinking and give them guidance to make it more expert-like. After a few years in that environment they turn into experts, not because there is something magic in the air in the research lab but because they are engaged in exactly the cognitive processes that are required for developing expert competence.

Once I realized this, I started to think how these ideas could be used to improve the teaching of undergraduate science. Of course it would be very effective to put every student into a research lab to work one-on-one with a faculty member rather than taking classes. While that would probably work very well and is not so different from my own education, obviously it is not practical as a widespread solution. So if the economic realities dictate that we have to use courses and classrooms, how can we use these ideas to improve classroom teaching? The key is to get these desirable cognitive activities, as revealed by research, into normal course activities.

I am not alone in coming to this conclusion. There is a significant community of science-education researchers, particularly in physics, who are taking this approach to the development and testing of new pedagogical approaches. This is paying off in clear demonstrations of improved learning. Indeed, some innovative pedagogical strategies are sufficiently mature that they are being routinely replicated by other instructors with similar results.

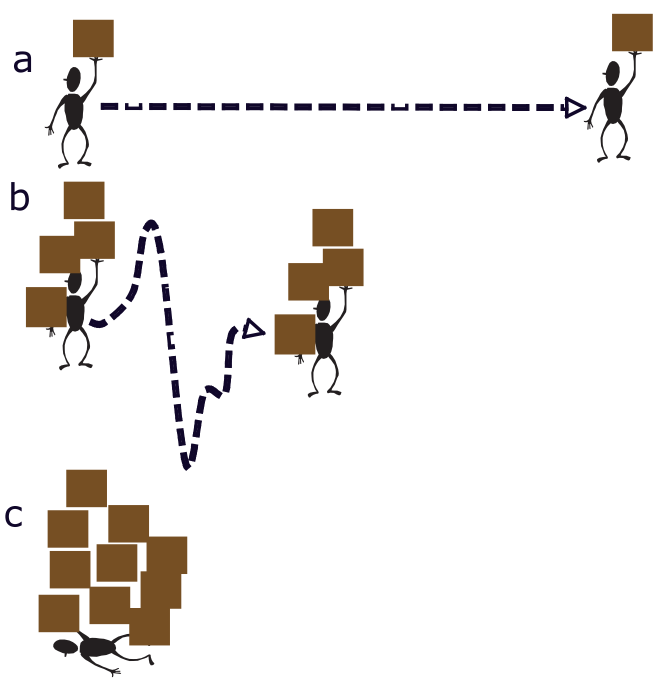
So what are a few examples of these strategies, and how do they reflect our increasing understanding of cognition?

Reducing Cognitive Load

The first way in which one can use research on learning to create better classroom practices addresses the limited capacity of the short-term working memory. Anything one can do to reduce cognitive load improves learning. The effective teacher recognizes that giving the students material to master is the mental equivalent of giving them packages to carry. (See Figure 5). With only one package, they can make a lot of progress in a hurry. If they are loaded down with many, they stagger around, have a lot more trouble, and can't get as far. And when they experience the mental equivalent of many packages dumped on them at once, they are squashed flat and can't learn anything.

So anything the teacher can do to reduce that cognitive load while presenting the material will help. Some ways to do so are obvious, such as slowing down. Others include having a clear, logical, explicit organization to the class (including making connections between different ideas presented and connections to things the students already know), using figures where appropriate rather than relying only on verbal descriptions and

Figure 5. Result of loading student up with low, medium, and high cognitive loads.



minimizing the use of technical jargon. All these things reduce unnecessary cognitive demands and result in more learning.

Addressing Beliefs

A second way teachers can improve instruction is by recognizing the importance of student beliefs about science. This is an area my own group studies. We see that the novice/expert-like beliefs are important in a variety of ways—for example they correlate with content learning and choice of major. However, our particular interest is how teaching practices affect student beliefs. Although this is a new area of research, we find that with rather minimal interventions, a teacher can avoid the regression mentioned above.

The particular intervention we have tried addresses student beliefs by explicitly discussing, for each topic covered, *why this topic is worth learning, how it operates in the real world, why it makes sense, and how it connects to things the student already knows*. Doing little more than this eliminates the usual significant decline and sometimes results in small improvements, as measured by our surveys. This intervention also improves student interest, because the beliefs measured are closely linked to that interest.

Stimulating and Guiding Thinking

My third example of how teaching and learning can be improved is by implementing the principle that effective teaching consists of engaging students, monitoring their thinking, and providing feedback. Given the reality that student-faculty interaction at most colleges and universities is going to be dominated by time together in the classroom, this means the teacher must make this happen first and foremost in the classroom.

To do this effectively, teachers must first know where the students are starting from in their thinking, so they can build on that foundation. Then they must find activities that ensure that the students actively think about and process the important ideas of the discipline. Finally, instructors must have mechanisms by which they can probe and then guide that thinking on an ongoing basis. This takes much more than just mastery of the topic—it requires, in the memorable words of Lee Shulman, "pedagogical content knowledge."

Getting students engaged and guiding their thinking in the classroom is just the beginning of true learning, however. This classroom experience has to be followed up with extended "effortful study," where the student spends considerably more time than is possible in the classroom developing expert-like thinking and skills.

Even the most thoughtful, dedicated teachers spend enormously more time worrying about their lectures than they do about their homework assignments, which I think is a mistake. Extended, highly focused mental processing is required to build those little proteins that make up the long-term memory. No matter what happens in the relatively brief period students spend in the classroom, there is not enough time to develop the long-term memory structures required for subject mastery.

To ensure that the necessary extended effort is made, and that it is productive, requires carefully designed homework assignments, grading policies, and feedback. As a practical matter, in a university environment with large classes the most effective way for students to get the feedback that will make their study time more productive and develop their metacognitive skills is through peer collaboration.

Using Technology

I believe that most reasonably good teachers could engage students and guide their thinking if they had only two or three students in the class. But the reality of the modern university is that we must find a way to accomplish this with a class of 200 students. There are a number of new technologies that, when used properly, can be quite effective at extending instructors' capabilities so that they can engage and guide far more students at once.

A caveat: Far too often, the technology drives instruction and student thinking rather than the educational purposes driving the development and use of the technology. A second caveat: There is far too little careful testing of various technologies' effectiveness in increasing the learning of real students. However, here are three demonstrably effective uses of technology.

"Just-in-time teaching" was introduced by Gregor Novack, Andy Gavrin, Evelyn Patterson, and Wolfgang Christian. The technique uses the Web to ask students questions concerning the material to be covered, questions that they must answer just before class. The students thus start the class already engaged, and the instructor, who has looked at the students' answers, already knows a reasonable amount about their difficulties with the topic to be covered.

A second technology that I have worked with extensively is personal-response systems or "clickers." Each student has a clicker with which to answer questions posed during class. A computer records each student's answer and can display a histogram of those responses. The clicker efficiently and quickly gets an answer from each student for which that student is accountable but which is anonymous to their peers.

I have found that these clickers can have a profound impact on the educational experience of students. The most productive use of clickers in my experience is to enhance the Peer Instruction (PI) technique developed by Eric Mazur, particularly for less active or assertive students.

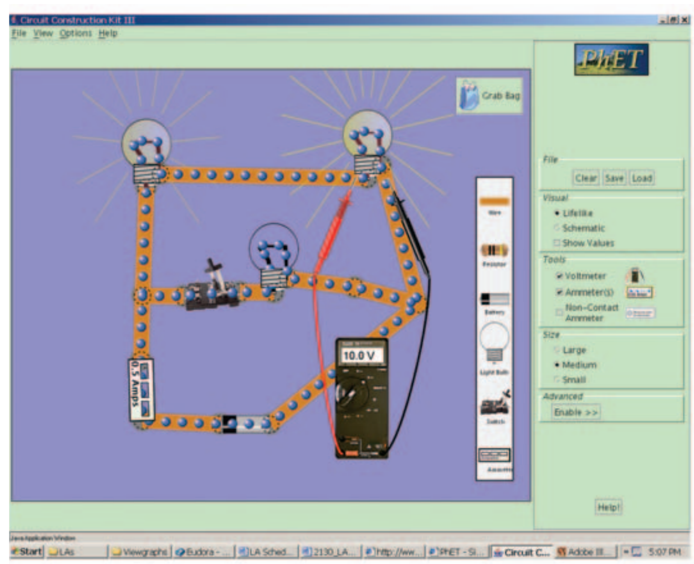
I assign students to groups the first day of class (typically three to four students in adjacent seats) and design each lecture around a series of seven to 10 clicker questions that cover the key learning goals for that day. The groups are told they must come to a consensus answer (entered with their clickers) and be prepared to offer reasons for their choice. It is in these peer discussions that most students do the primary processing of the new ideas and problem-solving approaches. The process of critiquing each other's ideas in order to arrive at a consensus also enormously improves both their ability to carry on scientific discourse and to test their own understanding.

Clickers also give valuable (albeit often painful) feedback to the instructor when they reveal, for example, that only 10 percent of the students understood what was just explained. But they also provide feedback in less obvious ways. By circulating through the classroom and listening in on the consensus-group discussions, I quickly learn which aspects of the topic confuse students and can then target those points in the follow-up discussion. Perhaps even more important is the feedback provided to the students through the histograms and their own discussions. They become much more invested in their own learning. When using clickers and consensus groups, I have dramatically more substantive questions per class period—more students ask questions and the students represent a much broader distribution by ethnicity and gender—than when using the peer-instruction approach without clickers.

A third powerful educational technology is the sophisticated online interactive simulation. This technique can be highly effective and takes less time to incorporate into instruction than more traditional materials. My group has created and tested over 60 such simulations and made them available for free (www.phet.colorado.edu). We have explored their use in lecture and homework problems and as replacements for, or enhancements of, laboratories.

The "circuit construction kit" is a typical example of a simulation (See Figure 6). It allows one to build arbitrary circuits involving realistic-looking resistors, light bulbs (which light up), wires, batteries, and switches and get a correct rendition of voltages and currents. There are realistic volt and ammeters to measure circuit parameters. The simulation also shows cartoon-like electrons moving around the circuit in appropriate paths, with velocities proportional to current. We've found this simulation to be a dramatic help to students in understanding the basic concepts of electric current and voltage, when substituted for an equivalent lab with real components.

Figure 6. Circuit constructions kit interactive simulation.



Source: Physics Ecducation Technology Project, University of Colorado

As with all good educational technology, the effectiveness of good simulations comes from the fact that their design is governed by research on how people learn, and the simulations are carefully tested to ensure they achieve the desired learning. They can enhance the ability of a good instructor to portray how experts think when they see a real-life situation and provide an environment in which a student can learn by observing and exploring. The power of a simulation is that these explorations can be carefully constrained, and what the student sees can be suitably enhanced to facilitate the desired

Table 1. Comparison of Learning Results from Traditionally Taught Courses and Courses Using Research-Based Pedagogy

Traditional Instruction	Research-Based Instruction
Retention of information from lecture: 10% after 15 minutes	Retention of information from lecture: More than 90 % after 2 days
Gain in conceptual understanding: 25%	Gain in conceptual understanding: 50-70%
Beliefs about physics and problem-solving: significant drop	A small improvement

learning. Using these various effective pedagogical strategies, my group and many others have seen dramatic improvements in learning.

Institutional Change

We now have good data showing that traditional approaches to teaching science are not successful for a large proportion of our students, and we have a few research-based approaches that achieve much better learning. The scientific approach to science teaching works, but how do we make this the norm for every teacher in every classroom, rather than just a set of experimental projects? This has been my primary focus for the past several years.

A necessary condition for changing college education is changing the teaching of science at the major research universities, because they set the norms that pervade the education system regarding how science is taught and what it means to “learn” science. These departments produce most of the college teachers who then go on to teach science to the majority of college students, including future school teachers. So we must start by changing the practices of those departments.

There are several major challenges to modifying how they educate their students. First, in universities there is generally no connection between the incentives in the system and student learning. A lot of people would say that this is because research universities and their faculty don’t care about teaching or student learning. I don’t think that’s true—many instructors care a great deal. The real problem is that we have almost no authentic assessments of what students actually learn, so it is impossible to broadly measure that learning and hence impossible to connect it to resources and incentives. We do have student evaluations of instructors, but these are primarily popularity contests and not measures of learning.

The second challenge is that while we know how to develop the necessary tools for assessing student learning in a practical, widespread way at the university level, carrying this out would require a significant investment. Introducing effective research-based teaching in all college science courses—by, for instance, developing and testing pedagogically effective materials, supporting technology, and providing for faculty development—would also require resources. But the budget for R&D and the implementation of improved educational methods at most universities is essentially zero. More generally, there is not the political will on campus to take the steps required to bring about cultural change in organizations like science departments.

Our society faces both a demand for improved science education and exciting opportunities for meeting those demands. Taking a more scholarly approach to education—that is, utilizing research on how the brain learns, carrying out careful research on what students are learning, and adjusting our instructional practices accordingly—has great promise. Research clearly shows the failures of traditional methods and the superiority of some new approaches for most students. However, it remains a challenge to insert into every college and university classroom these pedagogical approaches and a mindset that teaching should be pursued with the same rigorous standards of scholarship as scientific research.

Although I am reluctant to offer simple solutions for such a complex problem, perhaps the most effective first step will be to provide sufficient carrots and sticks to convince the faculty members within each department or program to come to a consensus as to their desired learning outcomes at each level (course, program, etc.) and to create rigorous means to measure the actual outcomes. These learning outcomes cannot be vague generalities but rather should be the specific things they want students to be able to do that demonstrate the desired capabilities and mastery and hence can be measured in a relatively straightforward fashion. The methods and instruments for assessing the outcomes must meet certain objective standards of rigor and also be collectively agreed upon and used in a consistent manner, as is done in scientific research.

Resources

- W. Adams et al. (2005), *Proceedings of the 2004 Physics Education Research Conference*, J. Marx, P, Heron, S. Franklin, eds., American Institute of Physics, Melville, NY, p. 45.
- R. Hake (1998), *The American Journal of Physics*. 66, 64.
- D. Hammer (1997), *Cognition and Instruction*. 15, 485.
- D. Hestenes, M. Wells, G. Swackhammer (1992), *The Physics Teacher*. 30, 141.
- Z. Hrepic, D. Zollman, N. Rebello. “Comparing students’ and experts’ understanding of the content of a lecture,” to be published in *Journal of Science Education and Technology*. A pre-print is available at http://web.phys.ksu.edu/papers/2006/Hrepic_comparing.pdf
- E. Mazur (1997), *Peer Instructions: A User’s Manual*, Prentice Hall, Upper Saddle River, NJ.
- G. Novak, E. Patterson, A.Gavrin, and W. Christian (1999), *Just-in-Time Teaching: Blending Active Learning with Web Technology*, Prentice Hall, Upper Saddle River, NJ.
- K. Perkins et al. (2005), *Proceedings of the 2004 Physics Education Research Conference*, J. Marx, P. Heron, S. Franklin, eds., American Institute of Physics, Melville, NY, p. 61.
- E. Redish (2003), *Teaching Physics with the Physics Suite*, Wiley, Hoboken, NJ.