THE NATIONAL ACADEMIES REPORT IN BRIEF

BIO 2010: Transforming Undergraduate Education for Future Research Biologists

Research in biology has undergone a major transformation in the last 10 to 15 years. Three powerful innovations – recombinant DNA, new instrumentation and the digital revolution – have combined to make biomedical research more quantitative and more closely connected to concepts in the physical, mathematical and information sciences. Researchers who once dedicated their lives to the study of a single gene, can now use sophisticated instrumentation and computer analysis to study the complex interactions of the more than 30,000 genes that make up the human genome.

In contrast, undergraduate biology education is still geared to the biology of the past. Although most colleges and universities require biology majors to enroll in courses in math, chemistry and physics, these subjects are not well integrated into biology courses. Furthermore, most courses, especially those for first-year students, are still primarily lecture-based, and do not convey the exciting reality of biology today.

What qualifications should a graduating biology major possess? What are the fundamental concepts of mathematics, chemistry, physics, computer science and engineering that will assist students in making interdisciplinary connections? How can universities implement new programs and what institutional barriers must be overcome?

The National Academies' report, *Bio2010: Transforming Undergraduate Education for Future Research Biologists*, identifies potential changes in undergraduate education designed to improve the preparation of students in the life sciences, with a particular emphasis on the education needed for future careers in biomedical research. The report looks at content, teaching approaches, curriculum requirements, funding and other issues.

Biology in Context: An Interdisciplinary Curriculum

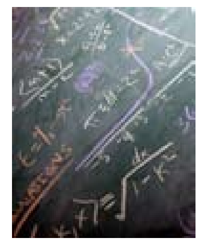
The modern biologist uses a wide array of advanced techniques, such as measuring instruments, novel imaging systems, computer analysis, and modeling that are rooted in the physical and information sciences. Focused laser beams allow manipulations of single molecules. X-ray sources are used to determine three-dimensional structures of proteins. Functional magnetic resonance imagers map activated regions of the brain. Computers now play a central role in the acquisition, storage, analysis, interpretation and visualization of vast quantities of biological data.

Understanding and applying these techniques requires access to a broader range of concepts and skill than past generations, much of it outside the traditional realm of biology education. Numerous studies and workshops have addressed the growing body of research at the intersection of biology with other disciplines, further supporting the need for more interdisciplinary education. Already, multidisciplinary projects are emphasized in solicitations for research grants.

The *Bio2010* report provides a consensus list of the central concepts of biology, chemistry, physics, math and computer science, and engineering that life science students should master in order to make novel interdisciplinary connections to address the reality of research today.

Central Concepts in Biology. Knowledge of diverse genomes, from bacteria to worms to flies to humans, is revealing recurring motifs and mechanisms and strengthening our appreciation for the fundamental unity of life. Variations on this unity lead to the extraordinary diversity of individual organisms. To understand this unity and diversity, teaching of biology students should focus on several central themes in multiple contexts. For example, the central theme of equilibria could be taught in a variety of contexts:

Living systems are far from equilibrium. They utilize energy, largely derived from photosynthesis, which is stored in high-energy bonds or ionic concentration gradients. The release of this energy is coupled to thermodynamically unfavorable reactions to drive biological processes.



Central Concepts in Math and Computer Science. The elucidation of the human genome has opened new vistas and highlighted the increasing importance of mathematics and computer science in biology. The current intense interest in genetic, metabolic and neural networks reflects the need of biologists to view and understand the coordinated activities of large numbers of components of the complex systems underlying life.

It is essential that biology undergraduates become quantitatively literate, studying the mathematical concepts of change, modeling, equilibria and stability, structure of a system, interactions among components, data and measurement, visualization, and algorithms. Every student should acquire the ability to analyze issues in these contexts in some depth, using analytical methods (e.g., pencil and paper) and appropriate computational tools. An appropriate course of study would include aspects of probability, statistics, discrete models, linear algebra, calculus and differential equations, modeling and programming.

Though all of these topics are offered in most universities and colleges, it is difficult for life science students to master the most essential concepts without taking a larger number of courses than can be accommodated in a biology major. The report recommends the creation of new courses that will cover the most relevant math concepts in less time in the context of biological problems.

As a good example, the University of Tennessee offers a two-semester course designed for life science majors that replaces the traditional calculus course (see Box 1). It introduces topics such as the mathematics of discrete variables, linear algebra, statistics, programming and modeling as applied to biological problems.

Box 1: Teaching that Works
Quantitative Life Sciences Education
at the University of Tennessee

This course sequence, developed by Dr. Louis Gross, provides an introduction to a variety of mathematical topics of use in analyzing problems arising in the biological sciences. The goal of the course is to show how mathematical ideas such as linear algebra, statistics and modeling can provide answers to key biological problems and to provide experience using computer software to analyze data and investigate mathematical models. Students are encouraged to formulate hypotheses that test the investigation of real world biological problems through the use of data.

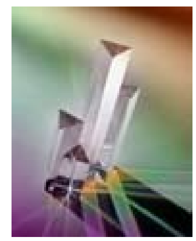
Each class session begins with students generating one or more hypotheses regarding a biological or mathematical topic germane to that day's material. For example, students go outdoors to collect leaf size data; they are then asked, Are leaf width and length related? Is the relationship the same for all tree species? What affects leaf size? Why do some trees have larger leaves than others? Each of these questions can generate many hypotheses, which students can evaluate after analyzing their data.

The program makes extensive use of graduate students in Tennessee's mathematical and computational ecology program because they are well positioned to explain the connections between mathematics and biology. More information on a quantitative curriculum for life science students can be found at www.tiem.utk.edu/~gross/quant.lifesci.html.

Central Concepts in Chemistry. Chemistry has always been an important sister science to biology, biochemistry, and medicine. Today, modern molecular and cell biology focuses on understanding the chemistry of genes and of cell structure. In the applied area, chemistry is central to modern agriculture, and biomedical engineering draws on chemistry for new materials. A thorough grounding in general and organic chemistry has historically required four semesters of chemistry courses, but could require fewer following an integrated restructuring.

The report recommends that biology majors receive a thorough education in chemistry, including aspects of organic, physical and analytical chemistry as well as biochemistry incorporated into new courses. Biology faculty could work in concert with chemistry colleagues to design curricula that will not only foster growth for aspiring chemists but also stimulate biology majors and those majoring in other disciplines. Core concepts include atoms, molecules, aqueous solutions, chemical reactions, energetics and equilibria, reaction kinetics, biomolecules, and materials.

Central Concepts in Physics. There is a set of basic physics concepts on which an understanding of biology can be built and that can be of aid in using increasingly sophisticated instrumentation. The typical calculus-based introductory physics course, which allocates a major block of time to electromagnetic theory and to many details of classical mechanics, is often the only option for biology students. The course emphasizes exactly solvable problems rather than the kinds of problems common in the life sciences. Illustrations involving modern biology are rarely given, and computer simulations are usually absent.



The report provides a list of physics concepts that life science majors should master including motion, dynamics and force laws; conservation laws and global constraints; thermal processes at the molecular level; waves, light, optics and imaging; and collective behavior and systems far from equilibrium. A redesigned physics course focused on these concepts would help biology students see how physicists think and how physics informs biology.

Central Concepts in Engineering. Biology increasingly involves the analysis of complex systems. Organisms can be analyzed in terms of subsystems having particular functions. Concepts in engineering can help biology students more easily describe and model how system functions result from constituent elements (see Box 2). For example, an effort to understand the locomotion of insects might be preceded by a laboratory involving an analysis of a simple legged robot, which provides a concrete model of the relation between the laws of physics and the problem of controlling directed movements.

The report recommends that life science majors be exposed to engineering principles and analysis that could include topics such as:

- the blood circulatory system and its control; fluid dynamics; pressure and force balance.
- material properties of biological systems and how structure relates to their function (e.g., wood, hair, cells).

Box 2: Teaching That Works On The Mechanics of Organisms

An upper-level course developed by Mimi Koehl at the University of California, Berkeley, brings biology and engineering together. It teaches functional morphology (how things move) in terms of mechanical design principles. Organisms are introduced as “Living Machines” and their abilities to fly, swim, parachute, glide, walk, run, buckle, twist and stretch are evaluated in the context of physics and engineering principles.

Students learn about the different types of fluid flow, the fluid dynamic forces of drag and lift, and how organisms live on wave-swept shores. They consider how mechanical properties change during the life of an organism, and the physics of shape change in morphogenesis, among other topics.

Energizing the Curriculum: New Content and Approaches

Successful interdisciplinary teaching will require both new materials and approaches. The need for teaching materials that will inform, enlighten and empower the next generation of researchers is crucial. New course designs and materials that encompass the highly interdisciplinary character of biology can accelerate the learning process and enable students to exercise their talents earlier in their careers.

An increasing number of today's college faculty recognize the significance of incorporating inquiry-based teaching and learning into their courses. The approach helps students to learn in the same way that scientists learn through research. Scientists ask questions, make observations, take measurements, analyze data, and repeat this process in an attempt to integrate new information. Teachers can use the approach in the classroom, labs, and the field.

The report presents several examples of ways to integrate two or more sciences together into one course as well as innovative teaching approaches that help communicate the excitement of science.

Modules for Course Enrichment

A logical first step in providing interdisciplinary course material is to use modules. The use of biological examples as modules in courses on chemistry, physics, computer science, and mathematics could help make those courses more relevant to future biological research scientists. Well-chosen examples that vividly present the biological pertinence of the physical or mathematical concepts under study can help students connect material taught in different courses.

A module can be presented in a single lecture or laboratory session, or over several sessions (see Box 3). Adaptable modules for course enrichment that take full advantage of interactive computer programs and multimedia educational tools are a very attractive complementary means of strengthening undergraduate biology education. Modules have been developed and integrated into science curricula with success at some institutions, but this approach has not been widely adopted at a majority of institutions nationwide.

Multiple independent groups have published modules or resources that can be used to enhance the teaching of undergraduate biology students. One group that has developed numerous modules for biology courses and laboratories is the BioQUEST Curriculum Consortium. The BioQUEST library is a peer-reviewed publication of computer-based curricular materials for biology education. The current volume contains more than 75 software simulations and supporting materials from diverse areas of biology.

Box 3: Teaching that Works:
The "Flu Module" at Carleton College

In his organic chemistry course, Dr. Jerry Mohrig introduced a "Flu Module" as a capstone, with a question that informs and drives the course. The capstone he presented was "Why do we get the flu every year?" Because a lot is known about the viral system, this capstone provides a modern, familiar context in which students can learn the basic chemistry of carbohydrates, proteins, molecular recognition, and cell-cell interactions. The module has been so successful, it is now used as a cohesive storyline every year.

Although most second-term organic chemistry courses include the basics of carbohydrate and amino acid chemistry, most students would be hard pressed to recognize or appreciate the great importance that carbohydrates have in biochemical recognition. The flu module focuses on how the interaction of carbohydrates and amino acids allow viral invasion of cells and also how therapeutic agents can be developed. Students are able to relate complex organic molecules to biological questions and they develop the confidence to do so.

Since he has been teaching the flu module, Dr. Mohrig has seen a significant increase in the interest in organic chemistry from the many biology students in the course.

The report offers ideas for potential modules, including:

What determines whether an epidemic waxes or wanes? In a simple model, a population consists of susceptibles who can contract a disease, infectives who can transmit it, and removals who have had the disease and are neither susceptible nor infective. Given an infection rate, a removal rate, and initial sizes of the three groups, one can calculate how the population evolves.

How do leopards get their spots and zebras get their stripes?
In 1952, Alan Turing published a seminal paper showing that an initially homogeneous distribution of chemicals can give rise to heterogeneous spatial patterns by reaction and diffusion.

Box 4: Teaching That Works
First-Year Seminar on Plagues

In the University of Oregon's first-year seminar, Plagues: The Past, Present, and Future of Infectious Diseases, professor Dan Udovic helps communicate the excitement of science. The course examines diseases such as malaria, bubonic plague, smallpox, polio, measles, and AIDS. In addition to the biology of the diseases, it also addresses their effects on populations and the course of history. Students investigate the conditions that influence the rate of spread of contagious diseases, and ways to prevent it. They discuss a number of ethical issues that arise in treating the sick, as well as development of policies intended to halt epidemics.

One segment of the course uses readings, discussions, computer modeling and lab activities to help students understand: (1) how the immune system works and why in some cases it doesn't; (2) why antibiotics work with some organisms but not others, and why many organisms are becoming resistant to antibiotics; (3) why so many new diseases seem to be suddenly appearing; (4) how vaccines work and why in some cases they don't; (5) how infectious diseases are transmitted; (6) why and how disease-causing organisms make humans sick; and (7) why most infectious diseases are usually not lethal.

Interdisciplinary Lectures and Seminars

In addition to modules, interdisciplinary lecture and seminar courses can give students a more realistic picture of how the sciences fit together. The report recommends that such courses be made available to students starting in their first year. At one end of the spectrum could be a first-year seminar with relatively few details and no prerequisites designed to "whet the appetite" of students who may or may not be majoring in biology. One excellent example is a first-year seminar on plagues that draws on disciplines outside the sciences (see Box 4).

At the other end of the spectrum is a capstone course for seniors with extensive prerequisites such as the "Mechanics of Organisms" course described in Box 2. At intermediate levels, a variety of course plans could incorporate material from the physical sciences and the underlying mathematical concepts and skills. A possible example is a course in quantitative physiology that explores blood circulation, gas exchange in the lung, control of cell volume, electrical activity of neurons and muscle mechanics.

Building on Concepts Through Laboratories

Laboratories can illustrate and build on the concepts covered in the classroom. Some concepts – such as error analysis, uncertainty, fluctuations and noise – are best learned through laboratory experiences. Once students have time to examine the specimens, materials, and equipment described in class, they are better prepared to carry out experiments. Project based laboratory work helps to stimulate student interest and participation, and is a choice arena to develop scientific writing, speaking, and presentation skills.

Interdisciplinary laboratories are a promising means of strengthening the physical sciences and quantitative background of life sciences majors and of introducing biology to students majoring in other fields. Harvey Mudd College has developed an introductory lab course designed to help students understand the research approach in science and the natural relationship between biology and other sciences (see Box 5).

The report proposes ideas for new labs in four disciplines: Physics, Engineering, Chemistry and Genomics, using a "crawl, walk, run" approach that helps students progress from step-by-step instructions to guidelines and examples, and finally to finding independent solutions to open-ended questions.

Box 5: Teaching that Works
Interdisciplinary Lab, Harvey Mudd College

In this team-taught course, students are led to understand the research approach in science. All experiments include technique development, instrumental experience, question formation and hypothesis testing, data and error analysis, oral and written reporting and most importantly, the opportunity to explore in an open-ended way details of phenomena that are familiar and of interest to students. Students are paired with a different partner for each experiment, developing teamwork skills in the process. Lab exercises include:

- Thermal properties of an ectothermic animal: Are lizards just cylinders with legs?
- Molecular weight of macromolecules: Is molecular weight always simple?
- Photosynthetic electron transport: How do biological systems convert physics into chemistry?

Incorporating Undergraduate Research

Many research scientists regard their undergraduate research experience as a turning point that led them to pursue research careers. By working as a partner in an active research group, undergraduates experience the rewards and frustrations of original research. Colleges and Universities should strive to make opportunities for independent research available to all students. They should regard the time faculty spend mentoring students one-on-one as teaching.

In spite of the overwhelming broad-based agreement that undergraduate research is good pedagogy, the educational value of undergraduate research for students and the impact of undergraduate research on faculty development as scholars and educators, has not been assessed in a systematic and intensive way. The report calls for further study on this important topic; assessment should be an integral part of the introduction of any new teaching approach.

Many schools have trouble finding the resources to offer independent research experiences to all students. A host of infrastructure limitations as well as an overwhelming number of biology students can combine to limit the number of students who can have opportunities for research experiences with independent work, at least early in an undergraduate career. One way to share the excitement of biology with students is to replicate the idea of independent work within the context of courses by incorporating inquiry-based learning, project labs, and group assignments. Although these methods have been used for ages, they can be "discovered" as new by successive generations of teachers and students.

MCAT: A Constraint on Curriculum Change

Innovation in undergraduate biology education is constrained by medical school admission requirements and specifically by the MCAT exam. The report recommends conducting an independent review of medical school admission requirements and testing in light of the rapidly changing nature of biological research, and the consequent need to transform undergraduate science education. A change in the MCAT itself, or in the way it is used for medical school admissions, would allow the biology curriculum to develop in a way that is beneficial to all students (including pre-med students) instead of allowing MCAT content to dictate what all students are taught.

Implementation: Building Momentum

Implementing the recommendations of this report will require a significant commitment of resources, both intellectual and financial. Successful redesign of courses and curricula requires a large investment of faculty time, departmental encouragement, and significant support from the college or university administration. Creation of new interdisciplinary majors is a significant challenge, often necessitating the hiring of new faculty with experience doing interdisciplinary research and teaching interdisciplinary topics.

Administrators need to recognize the time and effort required for change by encouraging faculty to take advantage of campus resources (such as teaching and learning centers and computer services) and supporting them for travel to conferences, workshops, and courses that will develop their teaching. Likewise, creation of new material will require the same commitment of funding and time. Potential formats of these needed teaching materials are diverse and complementary: printed books and guides, CDs and videos, Web sites, and interactive computer programs.

National Networks for Reform

Transformation of the undergraduate biology curriculum is tied to issues that extend beyond the reach of a single campus. Many individuals, institutions, organizations, and informal networks are working to address these issues. Significant change will require cooperation between these diverse groups.

Several disciplinary societies have education committees that address undergraduate teaching. Some, such as the American Society for Microbiology (ASM) and the American Institute of Biological Sciences (AIBS), employ full-time staff to make these efforts more successful. Another national group, Project Kaleidoscope (PKAL), has worked since 1989 to identify and disseminate sound principles and methods on which to base undergraduate education in the natural sciences and mathematics. Its members are faculty from all types of colleges and universities and all disciplines of the sciences. An important feature of PKAL is that participants in disciplinary and interdisciplinary workshops leave with specific action plans to implement on their home campus. It operates by looking for "what works" and encouraging others to apply those approaches in their own teaching.

Sources of Financial Support

Two principal organizations that have funded undergraduate biology education are National Science Foundation (NSF) and the Howard Hughes Medical Institute (HHMI). NSF supports a diverse array of projects in undergraduate science education. These projects fund activities such as research by undergraduates and development of teaching resources. HHMI invested more than $476 million between 1987 and 2001 to support improvements in biology education at 232 colleges and universities (HHMI Annual Report, 2001). Their investment has transformed biology instruction at these institutions, in ways ranging from developing new curricula, hiring new faculty, promoting faculty development, and supporting independent research by undergraduate students. Another private organization, the Whitaker Foundation, has spent considerable time and money on programs that enhance research and education in biomedical engineering.

The Central Role of Faculty Development: A Proposed Summer Institute

Undergraduate biology education can be effectively transformed only through close and sustained collaboration between colleges, universities, government agencies, professional societies, and foundations. It is often assumed that once a useful pedagogical approach is identified, it will be reproducible, easy to disseminate, and simple for another faculty member to implement in his/her home institution. The reality is that in teaching, as in research, faculty need to be trained to carry out new tasks and their efforts to do so need to be recognized.

The report proposes the creation of an annual summer institute dedicated to faculty development for biology professors (and other science faculty as appropriate) as an effective and appropriate means of building on the ideas of Bio2010 and fostering continued innovation in biology education.

The summer institute for biology education would be a venue for faculty to share information and experiences. It would help to increase communication between research universities and primarily undergraduate institutions by bringing faculty from both types of institutions together to learn from each other. It would facilitate the development, adaptation, and dissemination of innovative courses and course materials while providing training workshops for faculty and encouraging the development of a community of scientists/educators.

Potential topics include:

- The integration of quantitative examples into biology courses.
- Presenting examples of recent biological research that relies upon basic principles of chemistry or physics to undergraduate students.
- Ideas for exposing large numbers of students to research (how to think like a scientist): from laboratory courses to computer simulations to conceptual experiments.
- Developing teaching materials for the sharing of innovative approaches.
- Incorporating emerging research on cognition and assessment (See the 1999 NRC report *How People Learn* and the 2001 NRC report *Knowing What Students Know*).

A successful institute would require a partnership among a variety of institutions and organizations. A collaboration between the NAS, NRC, HHMI, and NSF would help to anchor the effort in the research establishment.

EXPERT CONSENSUS REPORT

For More Information: Contact Kerry Brenner of the National Academies' Board on Life Sciences at (202) 334-1245, kbrenner@nas.edu. **Bio2010: Transforming Undergraduate Education For Future Research Biologists** is available from the National Academy Press; 2101 Constitution Avenue, N.W, Washington, DC 20055; 800-624-6242 or 202-334-3313 (in the Washington metropolitan area); Internet: http://www.nap.edu.

Committee on Undergraduate Biology Education to Prepare Research Scientists for the 21st Century: **Lubert Stryer**, *Chair,* Stanford University; **Ronald Breslow**, Columbia University; **James Gentile**, Hope College; **David M. Hillis**, University of Texas, Austin; **John Hopfield**, Princeton University; **Nancy Kopell**, Boston University; **Sharon Long**, Stanford University; **Edward Penhoet**, Gordon and Betty Moore Foundation; **Joan Steitz**, Yale University; **Charles Stevens,** The Salk Institute for Biological Studies; **Samuel Ward**, University of Arizona; and **Kerry Brenner** (Study Director), The National Academies' Board on Life Sciences.

This study was supported by Contract Number N01-OD-4-2139, Task Order 64 between the National Academies and the National Institutes of Health and Award Number 71200-500115 between the National Academies and the Howard Hughes Medical Institute. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the organizations or agencies that provided support for the project.

BRUCE ALBERTS

Restoring Science to Science Education

What is taught in schools today is a caricature of science. Young people need to be introduced to science's full riches.

I love biology, and nothing in my four decades as a professional biological scientist has given as much satisfaction as seeing that spark of passion for the subject ignited in a young person. So it should be no surprise that nothing frustrates me more than to see that spark extinguished by misguided educators and mind-numbing textbooks. As I write this article, I have just returned from a discussion with 7th-grade students in San Francisco, at which they described their year-long biology class that they found tedious and anything but inspiring. The course was structured around a textbook that was among those officially selected by the state of California two years ago, after an elaborate and expensive process that California repeats every eight years. The exploration of the wonderful world of living things should be a fascinating delight for students. But in California, as in so many other parts of the United States and the world, most students gain no sense of the excitement and power of science, because we adults have somehow let science education be reduced to the memorization of "science key terms."

How did this happen? And what can we do to recover from this tragic misuse of our young people's time and effort in school?

Part of the answer to the first question lies in the fact that producing and selling textbooks is a big business, and the prevailing market forces have invariably led to mediocrity. Twenty years ago, the situation was elegantly described in a book whose title says it all: *A Conspiracy of Good Intentions: America's Textbook Fiasco*. Sadly, the situation has not changed. Much of the problem lies in the simplistic ways in which these books are usually evaluated, stressing the coverage of science terms and computerized text analyses.

In response to the education standards movement of the 1990s, the 50 states set about establishing their own very different sets of detailed science education standards. Because of this heterogeneity, textbook companies are forced to waste great amounts of time and resources on producing books that can satisfy the needs of as many states as possible. Even before the standards movement made things worse, U.S. textbooks had become known around the world for being "an inch deep and a mile wide." The result today is what I call science education as mentioning.

Take for example my field of cell biology, where for grades 5 to 8, the National Science Education Standards produced by the National Academies in 1996 emphasized understanding the essence of cells as the fundamental units of life, rather than learning the technical names of cell parts. The California state standards, on the other hand, stress all of these names. As a result, the adopted textbook for 7th grade contains five pages with 12 cell parts highlighted as key terms: including endoplasmic reticulum, Golgi body, lysosomes, mitochondria, and ribosomes. Because this 700-page book is forced by the California state standards to cover much of biology in similar detail, there is not enough room to explain most of these cell parts. Thus, for example, for the highlighted word "endoplasmic reticulum," the book simply states that "The endoplasmic reticulum's passageways help form proteins and other materials. They also carry material throughout the cell." Why should memorizing these two sentences be of any interest or importance to a 12-year-old? And what if anything will even the best students remember a year later?

Another part of the answer to why the United States has let science education go badly astray is that it is much easier to test for science words than it is to test for science understanding. The new age of accountability in U.S. education has led to a massive increase in testing, and the individual states have generally selected simple, low-cost, multiple-choice tests that can be rapidly scored. Because these high-stakes tests drive teachers to teach to them, they are thereby defining what science education means in our schools. This is a great tragedy, inasmuch as it trivializes education for young people. For far too many of them, education appears to be a largely senseless initiation ritual that is imposed on them by adults.

Consider, for example, the following question that is offered in California as a sample item for its 5th-grade science test:

A scientist needs to take a picture of the well-ordered arrangements of the atoms and molecules within a substance. Which of the following instruments would be best for the scientist to use?

A. A laser light with holograph
B. A seismograph
C. An electron microscope
D. A stereoscope

There are two major problems with this question. The first is that there is no right answer; an electron microscope does not generally have the resolution to decipher the relative arrangement of atoms. But much more important to me is the fact that learning the names of the different machines that scientists use is neither interesting nor relevant to the education of 10-year-olds.

The following anecdote illustrates how far we have strayed from what should be the central purpose of education: empowering students to learn how to learn on their own. A scientist parent notices that her elementary school child has thus far not been exposed to any science in school. As a volunteer teacher, she begins a science lesson by giving the children samples of three different types of soil. Each child is told to use a magnifying glass to examine the soils and write down what they observe in each sample. She waits patiently, but the children are unwilling to write anything. Her probing reveals that after three years of schooling, the students are afraid to express their views because they don't know "the right answer."

In fact, we know that life is full of ambiguous situations and that as citizens and workers we will have to solve many problems to which there is no right answer. To quote former Motorola CEO Robert Galvin, "Memorized facts, which are the basis for most testing done in schools today, are of little use in an age in which information is doubling every two or three years. We have expert systems in computers and the Internet that can provide the facts we need when we need them. Our workforce needs to utilize facts to assist in developing solutions to problems."

Life is nothing like a quiz show. If we adults allow students to believe that we think being educated means knowing all of the right answers, is it any wonder that nearly half of U.S. middle- and high-school students are found to be disengaged from their schooling?

The four strands of science learning

Ten years after producing the National Science Education Standards, the National Academies convened a distinguished committee of scientists and science education experts to take a fresh look at science education, considering all that had been learned in the interim. In 2007, this group produced the valuable report *Taking Science to School: Learning and Teaching Science in Grades K-8*. This analysis proposes that students who are proficient in science be expected to:

- know, use, and interpret scientific explanations of the natural world;
- generate and evaluate scientific evidence and explanations;
- understand the nature and development of scientific knowledge; and
- participate productively in scientific practices and discourse.

These four strands of science education were judged in the report to be of equal importance. Yet what is taught in most schools today, from kindergarten through introductory college classes, focuses almost exclusively on only a portion of the first of the four strands: teaching students to know scientific explanations of the natural world. Adopting the agenda in *Taking Science to School* will therefore require an ambitious effort to redefine the term "science education."

The source of the problem is college. For the most part, those of us who are scientists have made a mess of science education. Scientists are deeply engaged in attempting to unscramble the puzzle of how the world works, and we are thrilled to read about each year's startling advances that increase our understanding of the universe that surrounds us. It seems that each new finding raises new questions to be answered, providing an endless frontier for the next generation of scientists to explore. We believe passionately in the power of science to create a better world, as well as in the critical importance for everyone in society of the values and attitudes that science demands of scientists: honesty, a reliance on evidence and logic to make judgments, a willingness to explore new ideas, and a skeptical attitude toward simple answers to complex problems. But very little of this is conveyed to students in our teaching.

It is college science, both because of its prestige and because it is the last science course that most adults will take, that defines science education for future teachers and parents. And yet, when my science colleagues in academia teach a first-year course to college students, most will at best attempt to cover only the first of the four strands of science proficiency recommended in the National Academies report. Any redefinition of science education at lower levels will therefore require a major change in the basic college courses in biology, chemistry, physics, and earth sciences. Each must add an emphasis on the other three strands: on enabling college students to generate and evaluate scientific evidence and explanations; to understand the nature and development of scientific knowledge; and to participate productively in scientific practices and discourse. This requires that students actively experience science as inquiry in their classes, being challenged to collect data and solve problems in the way that scientists do. They will also need to explore a few aspects of the subject in depth and be challenged to come up with some of their own explanations, rather than simply parroting back what they have been told in lectures or in textbooks.

A four-part recipe for action

As in science, strategy is everything when attempting to tackle a difficult problem. And redefining science education along the lines recommended in the Academies' *Taking Science to School* report will certainly be difficult. To be effective, we need focus, and I therefore propose the following four-part strategy. Much of what I say here about how to move forward is reflected in the new *Opportunity Equation* report from the Carnegie Institute for Advanced Study Commission on Mathematics and Science Education, on which I served.

1) ***Enlist the National Academies, in collaboration with the National Science Teachers Association and the American Association for the Advancement of Science, to develop a pared-down set of common core standards for science education that reflect the principles in*** **Taking Science to School**. We have learned a great deal since 1996 from the response to the standards movement, and the governors and the chief state school officers of a majority of states now recognize the enormous disadvantages of having 50 different state standards for science education. The federal government should provide incentives to the states to sign on to this common standards movement. For example, it can help link the core standards to an energetic, nationwide development of high-quality curricula, to online teacher education and professional development resources, and to the development and continual improvement of a research-based system of quality assessments and standards, as described below.

2) ***Initiate a high-profile effort to produce quality assessments that measure student learning of all four strands of science proficiency***. Poor tests are currently driving poor teaching and learning, and the development of much better tests at all levels, from elementary school through introductory college courses, is therefore an urgent and challenging matter. Our nation's leaders should make this a matter of national service, recruiting a group of the very best scientists and science assessment experts to work together over successive summers, as was done in the post-Sputnik era in the United States. At the K-12 level, two very different types of high-quality tests will need to be developed around the core standards: formative assessments that teachers can use to measure student progress, so as to adjust their teaching appropriately during the school year; and summative assessments that the states will use for accountability purposes. At the college level, I envision an effort to develop and disseminate quality questions to be given on the final exam in introductory science courses. These would be designed to test for an understanding of the last three strands of science proficiency in *Taking Science to School* and therefore be applicable to courses in a variety of scientific fields. Has the course enabled the students to understand "science as a way of knowing", and has it prepared them to use scientific processes and evidence as adults? The professors who teach these courses are scientists and should therefore care deeply about the answer.

3) ***Link the core science standards and their associated assessments to an intensive research program in selected school districts, so as to provide the "ground truth" needed***

for their continuous improvement. Education is much too complex to ever expect to get it permanently right. What is the effect of the use of these standards and assessments in actual schools? In what ways are they driving high-quality teaching and learning of science? How should they be revised and improved? Answers to these types of questions require collaborations between skilled researchers and teachers, and they are critical if we are to develop the science of education that our nation needs. The Strategic Education Research Partnership (SERP) is a nonprofit institution that resulted from two successive studies by the National Academies that addressed the question, why is research knowledge used effectively to improve health, agriculture, and transportation, but not education? Now in its fourth year, SERP has demonstrated how highly effective research can be produced when groups of academics and practitioners collaborate in real school settings, setting an example for the substantial research effort that is essential to continuously improve science education.

4) ***Work to strengthen to strengthen the human resources systems of states and school districts so as to recruit, retain, and deploy a corps of highly qualified science and math teachers***. We must improve teacher retention by making school districts more attractive places to work. Teachers must be treated as professionals and teacher leaders recruited to help incorporate the wisdom of outstanding teachers into school, school system, and state education practices and polices. Without such advice from a district's best teachers, continual improvement cycles are unlikely to be maintained. The United States should consider international models, such as Singapore's, that incorporate rotating groups of outstanding teachers into the highest levels of the education policymaking apparatus. We should also consider the possibility of recruiting outstanding Ph.D. scientists into state and district office, so as to readily connect our schools to national and local resources in the scientific and science education communities.

The broad goal for science education must be to provide students with the skills of problem solving, communication, and general thinking required to be effective workers and educated citizens in the 21st century. Business and industry need problem solvers throughout the enterprise, as witnessed by many studies. These same skills are also crucial to enable everyone to navigate the increasingly complex and noisy world that we live in. Thus, they are essential to empower the citizens in a democracy to make wise judgments for themselves and their communities, which they are required to do in the midst of a cacophony of voices striving to sway rather than enlighten them.

Recommended reading

B. Alberts, "Redefining Science Education," *Science* 323 (2009): 427.

Carnegie–Institute for Advanced Study Commission on Mathematics and Science Education, *The Opportunity Equation* (2009).

R. Marshall and M. Tucker, *Thinking for a Living: Education and the Wealth of Nations* (New York: Basic Books, 1992).

National Research Council, *National Science Education Standards* (Washington, DC: National Academy Press, 1996).

National Research Council, *Taking Science to School: Learning and Teaching Science in Grades K-8*, R. A. Duschl, H. A. Schweingruber, and A. W. Shouse, eds. (Washington, DC: National Academies Press, 2007).

National Research Council, *Ready, Set, Science!: Putting Research to Work in K-8 Science Classrooms*, Sarah Michaels, Andrew W. Shouse, and Heidi A. Schweingruber, eds. (Washington, DC: National Academies Press, 2007).

L. Steinberg, B. Brown, and S. Dornbusch, *Beyond the Classroom* (Cambridge, MA: Touchstone Books, 1997).

H. Tyson-Bernstein, *A Conspiracy of Good Intentions: America's Textbook Fiasco* (New York: Basic Books, 1988).

Bruce Alberts (bruce.alberts@ucsf.edu) is professor of biochemistry and biophysics at the University of California, San Francisco, and editor-in-chief of Science *magazine.*

SCIENCE EDUCATION

Changing the Culture of Science Education at Research Universities

Universities must better recognize, reward, and support the efforts of researchers who are also excellent and dedicated teachers.

W. A. Anderson,[1] U. Banerjee,[2] C. L. Drennan,[3] S. C. R. Elgin,[4] I. R. Epstein[5], J. Handelsman,[6] G. F. Hatfull,[7] R. Losick,[8]* D. K. O'Dowd,[9]* B. M. Olivera,[10] S. A. Strobel,[6] G. C. Walker,[3] I. M. Warner[11]

Professors have two primary charges: generate new knowledge and educate students. The reward systems at research universities heavily weight efforts of many professors toward research at the expense of teaching, particularly in disciplines supported extensively by extramural funding (*1*). Although education and lifelong learning skills are of utmost importance in our rapidly changing, technologically dependent world (*2*), teaching responsibilities in many STEM (science, technology, engineering, and math) disciplines have long had the derogatory label "teaching load" (*3*, *4*). Some institutions even award professors "teaching release" as an acknowledgment of their research accomplishments and success at raising outside research funds.

Some studies suggest little or no correlation between effective teaching, judged by student evaluations, and research, as measured by productivity and citations (*5*). But we contend that excellence in research and teaching need not be mutually exclusive but are instead intertwined and can interact synergistically to increase the effectiveness of both. The distinction between research and teaching is somewhat artificial; professors teach students how to learn from known sources in the classroom, but also how to create new knowledge in their research laboratories.

We are Howard Hughes Medical Institute (HHMI) professors, biomedical research scientists who receive support from HHMI for creating new programs that more effectively engage students in learning science. We represent a diversity of institutions, from well-endowed private universities to large and underfunded state universities. In our opinion, science education should not only provide broad content knowledge but also develop analytical thinking skills, offer understanding of the scientific research process, inspire curiosity, and be accessible to a diverse range of students. We should be preparing students for a lifetime of learning about science with

an understanding of its power and limitations. Evidence shows that approaches that accomplish these goals include active, engaging techniques; inquiry-based approaches; and research courses (*6*).

All of us have experienced the challenges of balancing teaching and research. Our ability to invest time and effort into improving undergraduate science education has been facilitated by extramural support and outside recognition provided by HHMI. How do we now help transform our research universities so that the teaching of science and scientific research are seen more broadly as equally valuable and mutually reinforcing?

Departmental and university cultures often do not adequately value, support, and reward effective pedagogy. Outstanding contributions to research are evaluated by standard measures (e.g., publications and grant support); are recognized globally as well as locally; and are rewarded within the university (e.g., with promotions or salary increases). Teaching, in contrast, is rarely judged and appreciated from the outside and often only minimally from within (*7*, *8*). To establish an academic culture that encourages science faculty to be equally committed to their teaching and research missions, universities must more broadly and effectively recognize, reward, and support the efforts of researchers who are also excellent and dedicated teachers.

Toward this end, we advocate seven initiatives (reflecting our views and not necessarily those of HHMI). Although many of these ideas are not new, the context in higher education has changed because of widespread concern about educating enough scientists and scientifically literate citizens (*9*) and because resources that enable change have improved markedly in recent years (*10*–*12*).

1. *Educate faculty about research on learning.* No scientist would engage in research without exploring previous work in the field, yet few university educators read education research. Universities can demonstrate that they value teaching by treating it as a scholarly activity, such as through faculty training in teaching that is predicated on evidence-based (*10*, *13*) approaches. Training should address education theory, tested practices, and methods to assess learning. Teachers should have time to experiment with new methods, identify strategies that they can implement effectively in specific settings, and take advantage of resources that enable translation of learning principles to teaching practice. These practices must include strategies to engage students in introductory courses, arguably the highest-impact change that could be made (*10*, *13*–*15*).

2. Create awards and named professorships that provide research support for outstanding teachers. Many universities recognize outstanding teachers with a special title or a modest monetary award. Campus-wide recognition should also include unrestricted funds, as is typical for named professorships, which make it feasible to sustain research activities while continuing to contribute to teaching excellence. Incorporating talks by these individuals into distinguished science lecture series is an opportunity to introduce innovative pedagogy. This may also attract a new donor population interested in sponsoring named professorships for faculty who have demonstrated excellence in the training of future scientists. In addition to campus-wide recognition, annual department-level awards for excellence in teaching could provide funds, allocated by the dean, to support

[1]Howard University, Washington, DC 20059, USA. [2]University of California, Los Angeles, Los Angeles, CA 90095, USA. [3]Massachusetts Institute of Technology, Cambridge, MA 02139, USA. [4]Washington University, St. Louis, MO 63130, USA. [5]Brandeis University, Waltham, MA 02453, USA. [6]Yale University, New Haven, CT 06520, USA. [7]University of Pittsburgh, Pittsburgh, PA 15260, USA. [8]Harvard University, Cambridge, MA 02138,USA. [9]University of California, Irvine, Irvine, CA 92697, USA. [10]University of Utah, Salt Lake City, UT 84112, USA. [11]Lousiana State University, Baton Rouge, LA 70803, USA.

*To whom correspondence should be addressed. E-mail: losick@mcb.harvard.edu or dkodowd@uci.edu

CREDIT: ADAPTED BY P. HUEY/SCIENCE

the scholarly activities of the recipient. This would not only help more faculty who have devoted significant effort to teaching maintain their research programs but also demonstrate to their colleagues that the effort required to achieve teaching excellence is valued. Named lecture series could bring professors from other universities who are distinguished as both research scientists and teachers to deliver a campus-wide lecture on pedagogy and a discipline-specific lecture on their research.

3. *Require excellence in teaching for promotion*. Formal criteria for tenure and promotion typically indicate that teaching and scholarship carry equal weight. The reality, however, is that most research-oriented universities promote faculty primarily on the basis of research achievements and ability to raise money from sources outside the university. Promotion that requires excellence in teaching would go a long way toward improving education. We need to reach agreement on broad goals of college science education and establish a rubric for evaluating the extent to which teachers are meeting these goals. We must identify the full range of teaching skills and strategies that might be used, describe best practices in the evaluation of teaching effectiveness (*16*, *17*) (particularly approaches that encourage rather than stifle diversity), and define how these might be used and prioritized during the promotion process.

4. *Create teaching discussion groups*. Teaching is often conducted out of sight of departmental colleagues. Even in large introductory classes that are taught by teams of instructors, members of the team are often absent from each other's presentations. To address this, both junior and senior faculty members should be brought together in small, peer teaching groups. Group members would attend each other's lectures and provide confidential critiques that highlight the most effective or innovative teaching strategies used and identify steps to increase effectiveness. Such peer support demonstrates that the department values, and shares responsibility for, good teaching. Group members are exposed to a variety of teaching strategies, some of which may positively affect their own practices. Annual meetings of the faculty at large, hosted by the dean, should routinely include discussion of innovative teaching strategies.

5. *Create cross-disciplinary programs in college-level learning*. Researchers are often left to fend for themselves in attempting to learn and implement best teaching practices and in evaluating how well students learn. Yet many research universities have unexploited resources that could be drawn upon to improve college-level learning. For example, many universities have Departments or Schools of Education, but only a few of those [e.g., (*18*, *19*)] include in their mission undergraduate-level learning or robust connections to, and collaborations with, faculty members in STEM departments. Such collaborations could spawn innovative programs for experimentation and evaluation of teaching practices in the sciences. Psychology Departments often have experts in cognitive science who would be valuable participants in such programs. Though extensive discussion of best teaching practices is beyond the scope of this piece, we refer readers, e.g., to (*10*, *13*, *20–23*), as well as the Supporting Online Material.

6. *Provide ongoing support for effective science teaching.* The National Academies Summer Institute has helped faculty from almost 100 research universities implement principles of scientific teaching (*24*). University-based teaching centers provide professional support to faculty for assessment across disciplines, as well as training teaching assistants. Some STEM programs explicitly include in their mission the support and improvement of STEM education [e.g., (*25*, *26*)]. There is no better way to teach science than to engage students in doing science (*27–29*). To provide such opportunities for large numbers of students demands ingenuity, a willingness to seek out and support mentors, and provision of lab and field facilities. Projects that can draw on student peer-mentoring deserve special attention as benefiting both mentor and mentee.

7. *Engage chairs, deans, and presidents*. The critical ingredient in creating a culture that values and promotes both teaching and science is leadership. Chairs of STEM departments, deans of schools, and presidents of universities must elevate the status of the teacher-scientist, communicate the importance they attach to effective teaching, and create and support programs that promote innovation in science education [e.g., (*30*)].

The issues we raise go beyond the sciences. Increasingly, it seems that parents, funders of higher education, and others are questioning the value of the education that research universities provide. The continued vitality of research universities requires that we foster a culture in which teaching and research are no longer seen as being in competition, but as mutually beneficial activities that support two equally important enterprises: generation of new knowledge and education of our students.

References and Notes

1. V. Savkar, J. Lokere, *Time to Decide: The Ambivalence of the World of Science Toward Education* (Nature Education, Cambridge, MA, 2010); http://bit.ly/blWaPr.
2. J. B. Labov, A. H. Reid, K. R. Yamamoto, *CBE Life Sci. Educ.* **9**, 10 (2010).
3. A. L. Johnson, *J. Chem. Educ.* **8**, 115 (1931).
4. L. V. Koos, *The Adjustment of Teaching Load in a University* (Government Printing Office, Washington, DC, 1919); www.archive.org/details/adjustmentofteac00koosuoft.
5. H. W. Marsh, J. Hattie, *J. Higher Educ.* **73**, 603 (2002).
6. R. A. Duschl, H. A. Schweingruber, A. W. Shouse, Eds., *Taking Science to School: Learning and Teaching Science in Grades K–8* (National Academies Press, Washington, DC, 2007).
7. R. Wilson, *Chron. High. Educ.* **57**(3), A1 (2010); http://chronicle.com/article/Why-Teaching-Is-Not-Priority/124301/.
8. M. C. Taylor, *Chron. High. Educ.* **56**(42), A22 (2010); http://chronicle.com/article/An-Academy-to-Support/123779/.
9. Committee on Prospering in the Global Economy of the 21st Century *et al.*, *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future* (National Academies Press, Washington, DC, 2007).
10. J. Handelsman, S. Miller, C. Pfund, *Scientific Teaching* (Freeman, New York, 2007).
11. Education Resources Information Center, www.eric.ed.gov/.
12. Multimedia Educational Resource for Learning Online, www.merlot.org/merlot/index.htm.
13. W. B. Wood, *Annu. Rev. Cell Dev. Biol.* **25**, 93 (2009).
14. J. Fairweather, *Linking Evidence and Promising Practices in Science, Technology, Engineering, and Mathematics (STEM) Undergraduate Education: A Status Report for The National Academies National Research Council Board of Science Education* (BOSE) (BOSE, Washington, DC, 2010); www7.nationalacademies.org/bose/Fairweather_Commissioned-Paper.pdf.
15. N. Lasry, E. Mazur, J. Watkins, *Am. J. Phys.* **76**, 1066 (2008).
16. R. B. Barr, J. Tagg, *Change* **27**, 13 (1995); http://ilte.ius.edu/pdf/BarrTagg.pdf.
17. M. W. Klymkowsky, K. Garvin-Doxas, M. Zeilik, *Cell Biol. Educ.* **2**, 155 (2003).
18. University of Maryland Physics Education Research Group, www.physics.umd.edu/perg/.
19. Physics Education Research at Colorado University—Boulder, www.colorado.edu/physics/EducationIssues/.
20. D. Bok, *Our Underachieving Colleges: A Candid Look at How Much Students Learn and Why They Should Be Learning More* (Princeton Univ. Press, Princeton, NJ, 2005).
21. R. R. Hake, *Latin Am. J. Phys. Edu.* **1**, 24 (2007).
22. M. A. McDaniel, A. A. Callender, in *Learning and Memory: A Comprehensive Reference*, J. Byrne *et al.*, Eds. (Elsevier, Oxford, 2008).
23. H. Pashler *et al.*, *Organizing Instruction and Study to Improve Student Learning* (NCER 2007-2004, National Center for Education Research, U.S. Department of Education, Washington, DC, 2007).
24. C. Pfund *et al.*, *Science* **324**, 470 (2009).
25. STEM at North Carolina State University, http://stem.ncsu.edu/.
26. College of Engineering, https://engineering.purdue.edu/Engr/Academics/Schools.
27. R. Taraban, R. L. Blanton, Eds., *Creating Effective Undergraduate Research Programs in Science* (Teachers College Press, New York, 2008).
28. E. Seymour, A.-B. Hunter, S. L. Laursen, T. DeAntoni, *Sci. Educ.* **88**, 493 (2004).
29. D. Lopatto, *Cell Biol. Educ.* **3**, 270 (2004).
30. C. Wieman, K. Perkins, S. Gilbert, *Change* (March-April 2010).

Supporting Online Material

www.sciencemag.org/cgi/content/full/331/6014/152/DC1

10.1126/science.1198280

For complete article, see: American Journal of Physics, 74(1), 31-39 (2006)

QUESTION DRIVEN INSTRUCTION: TEACHING SCIENCE (WELL) WITH AN AUDIENCE RESPONSE SYSTEM

Ian D. Beatty, William J. Leonard, William J. Gerace, and Robert J. Dufresne

Physics Education Research Group
Scientific Reasoning Research Institute & Department of Physics
University of Massachusetts
Amherst, Massachusetts, USA

INTRODUCTION

Educational use of audience response systems (ARSs), a.k.a. "classroom response systems," is exploding in high schools and universities. One vendor claims over a million of their system's keypads have been used, in all 50 U.S. states and 10 countries worldwide, in thousands of K-12 schools and hundreds of universities (eInstruction, 2005). Several universities are beginning centralized programs to introduce and coordinate response system use across campus. A fringe technology ten years ago, ARS are entering the mainstream.

ARS have the potential to radically alter the instructional dynamic of our classrooms and impact student learning. However, for an instructor to realize this potential requires much more than merely learning to operate the technology. Response systems are a tool, not a solution. Their benefits are not conferred automatically; *how* they are used matters tremendously. To be fully effective, their use must be integrated into a larger, coherent pedagogic approach.

As part of the UMass Physics Education Research Group (UMPERG), we have worked with response systems for over a decade. In 1993 we began using *Classtalk*, a groundbreaking "classroom communication system" by Better Education Inc. In 1994 we received a U.S. National Science Foundation (NSF) grant (DUE-9453881) to deploy, develop pedagogy for, and study the impact of *Classtalk* (Dufresne *et al.*, 1996). In 1998 we began *Assessing-to-Learn*, an NSF-funded project (ESI-9730438) to seed response systems in secondary school physics classrooms and help teachers develop suitable pedagogic skills and perspectives (Beatty, 2000; Feldman & Capobianco, 2003). In 1999 we brought EduCue *PRS* (since purchased by GTCO CalComp and renamed *InterWrite PRS*) to UMass and began its dissemination across campus. As a sequel to *Assessing-to-Learn*, we are beginning a five-year NSF-funded project (ESI-0456124) to research secondary school science teachers' learning of response system pedagogy. Based on twelve years of experience with ARS — teaching, researching, and mentoring — we have developed a comprehensive perspective on the effective use of such systems for the teaching of science at both the secondary school and university levels.

In this chapter we will introduce that perspective. We will not attempt to describe how response systems work, report our personal experiences using them, or discuss detailed

Chapter 7 of *Audience Response Systems in Higher Education: Applications and Cases*, edited by David A. Banks. Idea Group Inc., Hershey PA. ISBN 1-59140-947-0 (hardcover), 1-59140-948-9 (paperback), 1-59140-949-7 (e-book). Preprint.

Designing Questions

The criteria for an effective QDI question are quite different from those for exam, quiz, and homework questions, and questions for formative assessment use should be engineered with great care. Elsewhere, we detail a theoretical framework for designing questions (Beatty *et al.*, submitted). In this section we present some general principles and suggestions.

Every question should serve an explicit pedagogic purpose: a specific activity to induce in students' minds, not just a piece of topic mater to cover. For example:

- Drawing out students' background knowledge and beliefs on a topic;
- Making students aware of their own and others' perceptions and interpretations of a situation;
- Discovering particular confusions, misconceptions, and knowledge gaps;
- Distinguishing similar concepts;
- Realizing connections or similarities between different concepts;
- Elaborating the understanding of a concept; and
- Exploring the implications of an idea in a new or extended context.

Computational or simple factual questions, and those that probe memory rather than understanding and reasoning, are of little value. Questions that have students compare two situations, or make predictions and explore causal relationships, are particularly powerful. Good questions push students to reason qualitatively and draw conclusions from a conceptual model. If an instructor can anticipate likely misunderstandings and points of confusion, she should design questions to "catch" students in those, get them articulated, and resolve them through discussion.

Unlike exam questions, ARS questions for QDI benefit from ambiguity. An ambiguous feature sensitizes students to the feature's importance and implications, teaches them to pay attention to subtleties, and motivates discussion of what aspects of a question are important and how they matter. In this way, students can be led to contemplate not just one question but a family of related questions. Similarly, including irrelevant information or omitting necessary information can be beneficial, helping students learn to evaluate what information an answer requires. Questions need not be "fair" or even well defined, since we seek not to evaluate students but rather to help them learn to reason, think defensively, and answer future questions — especially the vague, fuzzy kind often encountered outside the classroom. (However, some questions should be straightforward and provide students with confirmation that they do in fact "get" a particular topic: this is useful feedback to them, and also good psychology.)

A question that elicits a spectrum of answers is generally more productive than one all students agree upon: it provides fodder for discussion and disagreement, leading to engagement and learning.

When designing sets of related or sequential questions, instructors should remember that students experience significant "cognitive load" when reading and interpreting a new scenario. Reusing a situation for multiple questions is efficient, allowing students to concentrate on the relevant aspects of the question at hand and realize the implications of features that do change. Conversely, asking questions with the same conceptual content set in completely different circumstances helps students learn to see through a situation's "surface features" to its "deep structure," and to distinguish the core principles from the details.

When and how a question is presented can shape the depth, quality, and character of resulting student thought and interaction. Students tend to assume that the question relates to whatever has recently transpired in the course, and will apply knowledge accordingly. This can lead to "pigeonhole" learning in which concepts are assimilated chronologically and only accessible within a narrow context, rather than being organized into an interlinked, versatile hierarchy. A careful instructor will mix questions of varying types and topics, and include integrative questions that connect recent ideas with earlier ones.

Classroom Management

Perhaps the most initially daunting (and ultimately exhilarating) aspect of QDI is the necessity of giving up control of the classroom. A lecture is predictable and controlled, with attention safely focused on the instructor. QDI, however, necessarily turns the classroom over to students for dialogue and debate. We must learn to manage the apparent chaos rather than attempting to rigidly control it. Furthermore, the principle of "agility" means we must be prepared — even eager — to modify or discard a lesson plan and extemporize.

Some basic attention-management techniques help considerably. For example, one challenge is to recapture students' attention after they have been discussing a formative assessment question among themselves. An ARS helps dramatically here: by collecting answers (with a time limit) and projecting the resulting histogram on a large screen, attention is redirected to the front of the classroom. Students are naturally curious about each other's answers. Another challenge we face is determining how much time to allow students for small-group discussion of a formative assessment question. Noise level is a clue: when a question is shown, the class is initially quiet as students read and digest it; the noise level then rises as they discuss the question, and begins to fall as they reach resolution. This is an appropriate time to collect answers, display the histogram, and begin the whole-class discussion.

Encouraging students to speak up during the whole-class discussion is crucial. When soliciting volunteers to argue for various answers, we should maintain a strict poker face and not divulge which answer (or answers) is (or are) correct (if any). Allow the students to challenge each other's arguments. If nobody will defend a particular position, ask if anyone else will speculate on the reasoning that might lead to such an answer. (Nothing motivates a student to speak up like having someone else misrepresent his position.) Paraphrasing a student's statements can be valuable, perhaps even necessary in an acoustically challenging room, but we must be careful to stay as close as possible to the student's vocabulary and check with the student that the paraphrase is satisfactory.

When we decide to drop our poker-face and offer a little illumination of our own, we should downplay notions of "correct" and "incorrect" lest we focus students' attention too much on getting the right answers rather than on reasoning and understanding. Instead of commenting that a particular answer or argument is wrong, we can often say "that would be correct *if*…", indicating some similar situation or question for which it would be valid. This is not only less disconfirming to the student and less deterring to others, it is also more pedagogically productive for all the reasons that "compare and contrast" questions are powerful. We have found that often, students who appear to be offering a wrong answer are instead offering the right answer to the wrong question. Unless they are sensitized to this, telling them they are simply incorrect is confusing rather than enlightening.

Moderating a whole-class discussion presents us with the great danger of making the class instructor-centered rather than student-centered. Working from within students' perceptions and arguments, rather than making assertions from authority, helps to avoid this. Similarly, if a question contains ambiguities or errors, allowing students to discover these or drawing them out during discussion is preferable to announcing corrections as the question is presented. We should strongly resist any temptation to read a presented question out loud or to talk while students are engaged in small-group dialogue and answering. If we seek active learning, we must give them space to do it!

Tactical Decisions: Modeling Students' Needs

Though managing the classroom may be the most daunting aspect of QDI, modeling a class-full of students and deciding how best to interact with them is the most enduringly difficult aspect, and it is the very heart of the approach. It requires two distinct skills: modeling and interacting with an individual student, and handling an ensemble of individuals in parallel. Neither comes easily, and both can be truly mastered only by repeatedly trying, occasionally missing the mark, reflecting, and trying again. However, we offer some general advice to help the interested instructor get started.

Interacting "agilely" with a student is a modeling process closely analogous to the scientific method: observe, form a model, make a prediction based on the model, test the prediction, refine the model, and iterate (Gerace, 1992). In this context, we want to model both the student's knowledge (especially the gaps) and her thinking processes (especially the weaker skills). In contrast to a traditional lecture, we must practice "active listening": listening carefully and patiently to what she says and how her responses, questions, and other behaviors vary from what we expect. Even when we think we know what she is in the process of asking, we should let her finish: both out of respect and because every nuance of her utterance is valuable data. We will often answer a question with a question, not just rhetorically but to understand better why the student needs to ask hers. Our goal is not to answer her question, but to understand why she needs to ask it.

Rather than concentrating on the knowledge we wish to communicate, a less direct approach is often more effective: trying to figure out what prevents her from understanding, and then attacking the obstacles. This sleuthing out of the roots of confusion is an iterative and thoughtful process on our part. Of course, a rich knowledge of pedagogic theory and common points of confusion are useful. If we find ourselves stumped trying to help an individual, other students in the class can assist. They can often understand their peers better than we.

Clearly, carrying out such an attention-demanding, thorough process with every student in a full-sized class is impossible. We must try to track an array of typical or likely student mentalities, test the class for the accuracy of this array, and teach to it. For example, if a formative assessment question elicits a range of answers, we can ascribe a putative explanation to each one for why a student might select it, and that becomes our working model. Since we have probably prepared the answer set in advance, we should already have ideas about why each answer might be chosen. The distribution of class answers "fits" the model to the class.

This approach does not attach a model to any specific individual in the class. A complementary approach is to identify certain students as representatives of various sub-populations within the class, and then build and maintain as rich a model as possible of each. This can be very powerful: it is easier for us to think in detail about a real, specific individual than an abstract entity, and yet students generally have enough in common that by addressing one student's needs, we impact many. As a side benefit, the more we treat students as three-dimensional individuals, pay real attention to them, and try to understand their thinking, the more they will believe we care about them personally and are "on their side," and the less adversarial the instructional dynamic will be.

Coaching

QDI requires students to adopt a role they might not be accustomed to from more traditional instruction. Our experience is that the vast majority of students express positive feelings about ARS use and QDI after they have adjusted to it, but this adjustment takes time, and some initially greet it with fear and resentment. Students habituated to success under traditional instruction are most likely to be hostile: they have "mastered the game," and now the rules are being changed. Others object out of simple laziness: they are being asked to engage in thought and activity during class, and that is effortful and at times frustrating. They are also expected to complete assignments beforehand so as to be prepared for class. Many are uncomfortable with the idea that they are accountable for material not directly presented in lecture. Inducing students to become participating, invested learners is vital to the success of QDI, and meta-communication is our most powerful tool for achieving that. We can explain to students why we are doing what we are doing, at both the immediate and strategic levels, and how students will benefit. We can talk frankly about the obstacles students will likely encounter and how they can most effectively surmount them. In other words, we can explicitly address learning and communication as part of the "course material."

Some student perceptions merit particular attention. Initially, students will probably view formative assessment questions as mini-tests to be passed or failed. If this attitude is allowed to persist, it will sour them on the formative assessment approach and prevent them from fully and constructively engaging in the process. We must explicitly discuss the purpose of formative assessment and stress that the point is not to answer correctly, but to discover previously unnoticed aspects of the subject and of their own understanding. We must consistently reinforce this position by deemphasizing the correctness of answers and emphasizing reasoning and alternative interpretations. Assigning course credit for "correct" answers is massively counterproductive.

Tactics for Questions FTEP Learning goals workshop, February 2006

Here are some common tactics that may help you write questions to assess learning goals. (Beatty, 2005)

1. Remove inessential details to focus students' attention where you want it.

2. Have students compare two things. Their attention will naturally be drawn to the differences between them.

3. Ask a familiar question about an unfamiliar situation to draw students' attention to the ways the new situation differs from a familiar one.

4. Ask a series of two questions. The first is a trap intended to make students commit a common error. Before reviewing the first question, ask a second which makes them aware of the error they have just committed. This technique can help them discover the mistake they made.

5. Require students to use different representations. Ask them to explain in words the meaning of a mathematical formula. Ask them to use information from a graph in a mathematical formula. Ask them to graph data in a table.

6. Present students with a set of processes or objects and ask them to determine subsets within the items presented.

7. Direct the strategy to force students to use more than one method. If students commonly solve a type of problem one way, require that they use a different method.

8. Include extraneous information or omit necessary information so that students think more carefully about what they need to solve the problem. If they are always provided with only the information needed, an important part of the problem solving has been done for them. "Not enough information is given" can be the correct answer for some questions.

Chapter 5

Learning

Elizabeth L. Bjork and Robert Bjork

Making Things Hard on Yourself, But in a Good Way: Creating Desirable Difficulties to Enhance Learning

Chapter in
Psychology + the Real World
by FABBS Foundation,
M.A. Gernsbacher,
RW Pew
L. M. Hough
et al.

Making Things Hard on Yourself, But in a Good Way: Creating Desirable Difficulties to Enhance Learning

Elizabeth Ligon Bjork and Robert A. Bjork
University of California, Los Angeles

Please describe your current position and research interests.

Elizabeth Ligon Bjork: I am Professor of Psychology and Immediate-Past Chair of the University of California, Los Angeles, Academic Senate. My research interests have included visual attention and developmental processes but now focus on practical and theoretical issues in human memory and learning, particularly the role that inhibitory processes play in an efficient memory system.

Robert A. Bjork: I am Distinguished Professor and Chair of Psychology at the University of California, Los Angeles. My research focuses on human learning and memory and on the implications of the science of learning for instruction and training.

How did you get interested in studying the facilitating effect of apparent impediments to learning?

Elizabeth Bjork: My interests in optimizing learning were triggered by interactions with students lamenting during office hours how hard they had studied, only then to perform poorly on a just-given exam. This motivated me to examine why students' study activities were sometimes so ineffective.

Robert Bjork: My interests go back to my efforts—as a graduate student—to understand the relationship of forgetting and learning, especially why inducing forgetting often enhances subsequent learning. My interests in the application of "desirable difficulties" were fanned by my experiences teaching and coaching and from what I learned as Chair of the National Research Council Committee on Techniques for the Enhancement of Human Performance (1988–1994).

What has been the real-world impact of this work?

Overall, the impact has been slight. There are multiple indications, however, that the impact of basic research findings on educational practices is increasing and that, in particular, optimizing instruction will require unintuitive innovations in how the conditions of instruction are structured.

As teachers—and learners—the two of us have had both a professional and personal interest in identifying the activities that make learning most effective and efficient. What we have discovered, broadly, across our careers in research, is that optimizing learning and instruction often requires going against one's intuitions, deviating from standard instructional practices, and managing one's own learning activities in new ways. Somewhat surprisingly, the trials and errors of everyday living and learning do not seem to result in the development of an accurate mental model of the self as learner or an appreciation of the activities that do and do not foster learning.

The basic problem learners confront is that we can easily be misled as to whether we are learning effectively and have or have not achieved a level of learning and comprehension that will support our subsequent access to information or skills we are trying to learn. We can be misled by our subjective impressions. Rereading a chapter a second time, for example, can provide a sense of familiarity or perceptual fluency that we interpret as understanding or comprehension, but may actually be a product of low-level perceptual priming. Similarly, information coming readily to mind can be interpreted as evidence of learning, but could instead be a product of cues that are present in the study situation, but that are unlikely to be present at a later time. We can also be misled by our current performance. Conditions of learning that make performance improve rapidly often fail to support long-term retention and transfer, whereas conditions that create challenges and slow the rate of apparent learning often optimize long-term retention and transfer.

Learning versus Performance

This apparent paradox is a new twist on an old and time-honored distinction in psychology—namely, the distinction between learning and performance. Performance is what we can observe and measure during instruction or training. Learning—that is, the more or less permanent change in knowledge or understanding that is the target of instruction—is something we must try to infer, and current performance can be a highly unreliable index of whether learning has occurred.

Learning Without Performance and Performance Without Learning

Decades ago, learning theorists were forced to distinguish between learning and performance because experiments revealed that considerable learning could happen across a period when no change was apparent in performance. In latent-learning experiments with animals, for example, periods of free exploration of a maze, during which the animal's behavior seemed aimless, were shown—once reward was introduced—to have produced considerable learning. Similarly, in research on motor skills, investigators found that learning continued across trials during which the build-up of fatigue suppressed performance.

More recently, a variety of experiments—some of which we summarize below—have demonstrated that the converse is true as well: Namely, substantial improvements in performance across practice or training sessions can occur without significant learning (as revealed after a delay or in another context). To the extent, therefore, that people interpret current performance as a valid measure of learning, they become susceptible to misassessing whether learning has or has not occurred.

Storage Strength Versus Retrieval Strength

At a theoretical level, we (Bjork & Bjork, 1992) distinguish between the storage strength and the retrieval strength of information or skills stored in memory. Storage strength reflects how entrenched or interassociated a memory representation is with related knowledge and skills, whereas retrieval strength reflects the current activation or accessibility of that representation and is heavily influenced by factors such as situational cues and recency of study or exposure. Importantly, we assume that current performance is entirely a function of current retrieval strength, but that storage strength acts to retard the loss (forgetting) and enhance the gain (relearning) of retrieval strength. The key idea for present purposes is that conditions that most rapidly increase retrieval strength differ from the conditions that maximize the gain of storage strength. In other words, if learners interpret current retrieval strength as storage strength, they become susceptible to preferring *poorer* conditions of learning to *better* conditions of learning.

Introducing Desirable Difficulties to Enhance Learning and Instruction

So what are these better conditions of learning that, while apparently creating difficulty, actually lead to more durable and flexible learning? Such desirable difficulties (Bjork, 1994) include varying the conditions of learning, rather than keeping them constant and predictable; interleaving instruction on separate topics, rather than grouping instruction by topic (called blocking); spacing, rather than massing, study sessions on a given topic; and using tests, rather than presentations, as study events.

Before proceeding further, we need to emphasize the importance of the word *desirable*. Many difficulties are undesirable during instruction and forever after. Desirable difficulties, versus the array of undesirable difficulties, are desirable because they trigger encoding and retrieval processes that support learning, comprehension, and remembering. If, however, the learner does not have the background knowledge or skills to respond to them successfully, they become undesirable difficulties.

Varying the Conditions of Practice

When instruction occurs under conditions that are constrained and predictable, learning tends to become contextualized. Material is easily retrieved in that context, but the learning does not support later performance if tested at a delay, in a different context, or both. In contrast, varying conditions of practice—even varying the environmental setting in which study sessions take place—can enhance recall on a later test. For example, studying the same material in two different rooms rather than twice in the same room leads to increased recall of that material (Smith, Glenberg, & Bjork, 1978)—an empirical result that flies in the face of the common how-to-study suggestion to find a quiet, convenient place and do all your studying there.

A study of children's learning provides a striking illustration of the benefits of varying conditions of practice. Eight-year-olds and 12-year-olds practiced throwing beanbags at a target on the floor with their vision occluded at the time of each throw. For each age group, half of the children did all their practicing throwing to a target at a fixed distance (for example, 3 feet for the 8-year-olds), while the other half threw to targets that were closer or farther away. After the learning sessions and a delay, all children were tested at the distance used in the fixed-practice condition for their age group (Kerr & Booth, 1978).

Common sense would suggest that the children who practiced at the tested distance would perform better than those who had never practiced at that distance, but the opposite was true for both age groups. The benefits of variation—perhaps learning something about adjusting the parameters of the motor program that corresponded to the throwing motion—outweighed any benefits of being tested at the practiced distance. Many other studies have shown that when testing after training takes place under novel conditions, the benefits of variation during learning are even larger.

Spacing Study or Practice Sessions

The effects of distributed practice on learning are complex. Although massing practice (for example, cramming for exams) supports short-term performance, spacing practice (for example, distributing presentations, study attempts, or training trials) supports long-term retention. The benefits of spacing on long-term retention, called the spacing effect, have been demonstrated for all manner of materials and tasks, types of learners (human and animal), and time scales; it is one of the most general and robust effects from across the entire history of experimental research on learning and memory.

Rather than describing any of the myriad studies that have demonstrated the benefits of spacing, we will simply stress the importance of incorporating spacing and avoiding massing in managing learning. Massing repeated-study activities is often not only convenient, but it can also seem logical from the standpoint of organizing one's learning of different topics, and it frequently results in rapid gains in apparent learning. Good test performance following an all-night cramming session is certainly rewarding, but little of what was recallable on the test will remain recallable over time. In contrast, a study schedule that spaces study sessions on a particular topic can produce *both* good exam performance *and* good long-term retention. Furthermore, because new learning depends on prior learning, spacing study sessions optimally can also enhance transfer of knowledge and provide a foundation for subsequent new learning.

Interleaving versus Blocking Instruction on Separate To-Be-Learned Tasks

Interleaving the practice of separate topics or tasks is an excellent way to introduce spacing and other learning dynamics. In a classic comparison of interleaving and blocking (Shea & Morgan, 1979), participants practiced three different movement patterns, each requiring the participants to knock down three of six hinged barriers rapidly on a pinball-like apparatus in a prescribed order. All participants received 18 trials on each pattern, but in the interleaved condition, practice on a given trial was randomly determined, whereas in the blocked condition, one pattern at a time was practiced.

As you probably suspect, participants given blocked practice improved more rapidly than those given interleaved/random practice. Thus, if the researchers had stopped their study at the end of training, blocking of practice would have seemed the superior learning procedure. But, instead, participants returned 10 days later and were retested under either blocked or interleaved/random conditions. Under interleaved/random testing conditions, participants who had practiced under interleaved conditions performed far better than did the blocked-practice participants, who appeared, when tested under a random schedule, to have learned virtually nothing. Under blocked testing conditions, performance was essentially the same for both groups, but the small difference still favored the interleaved group.

The skills literature includes many replications of the pattern that blocked practice *appears* optimal for learning, but interleaved practice actually results in superior long-term retention and transfer of skills, and research illustrates that learners—as well as instructors—are at risk of being fooled by that pattern. For example, when participants who had learned three different keystroke patterns were asked to predict their performance on a test the next day, those given interleaved practice predicted their performance quite closely, whereas those given blocked practice were markedly overconfident (Simon & Bjork, 2001). In effect, the blocked-practice group misinterpreted their good performance during practice as evidence of long-term learning, rather than a product of the local (that is, blocked) conditions. Said differently, they misinterpreted the retrieval strength of a given keystroke pattern as an index of its storage strength.

Other results illustrate that the benefits of interleaved practice extend beyond the learning of motor skills. For example, when participants were asked to learn formulas for calculating the volumes of different solids, such as a truncated cone, in either a blocked or interleaved manner, interleaved instruction enhanced performance on a delayed test. The size of the long-term advantage of interleaved practice was striking: 63 percent versus 20 percent of new problems worked correctly a week later (Rohrer & Taylor, 2007).

More recently and surprisingly, we have found that interleaving even enhances inductive learning (Kornell & Bjork, 2008). When participants were asked to learn the styles of each of 12 artists based on a sample of 6 paintings by each artist, interleaving a given artist's paintings among the paintings by other artists—versus presenting that artist's paintings one after another (blocking)—enhanced participants' later ability to identify the artist responsible for each of a series of new paintings. This result is surprising because blocking would seem to make it easier to note the commonalities that characterize a particular artist's style. Indeed, as illustrated in Figure 1, the majority of participants—when asked after the test whether interleaving or blocking had helped them learn an artist's style better—definitely had the impression that blocking had been more effective than interleaving, the op-

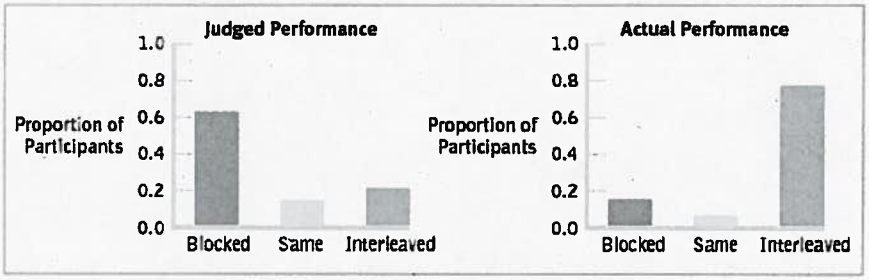


FIGURE 1 The left panel shows the proportion of participants who selected "blocked," "interleaved," or "the same" in response to the question: "Under which condition do you believe you learned better?" The right panel indicates the proportion of participants who actually performed better in the blocked or interleaved conditions or performed the same in the two conditions. (Kornell & Bjork, 2008)

posite of their actual learning. Blocking may indeed have facilitated noticing commonalities, but the final test required distinguishing among the artists, and interleaving may have fostered learning the differences as well as similarities among the styles of different artists.

Why might interleaving enhance long-term retention and transfer? One theory suggests that having to resolve the interference among the different things under study forces learners to notice similarities and differences among them, resulting in the encoding of higher-order representations, which then foster both retention and transfer. Another explanation suggests that interleaving forces learners to reload memories: If required to do A, then B, then C, and then A again, for example, the memory for how to do A must be reloaded a second time, whereas doing A and then A again does not involve the same kind of reloading. Such repeated reloadings are presumed to foster learning and transfer to the reloading that will be required when that knowledge or skill is needed at a later time.

From the standpoint of our theoretical framework (Bjork & Bjork, 1992), learning from reloading is an instance of a broader dynamic in human memory: Namely, that forgetting (losing retrieval strength) creates the opportunity for increasing the storage strength of to-be-learned information or skills. Said differently, when some skill or knowledge is maximally accessible from memory, little or no learning results from additional instruction or practice.

Generation Effects and Using Tests (Rather Than Presentations) as Learning Events

An effect that rivals the spacing effect for its generality and its significance for instruction and learning is the generation effect, which refers to the long-term benefit of generating an answer, solution, or procedure versus being presented that answer, solution, or procedure. Basically, any time that you, as a learner, look up an answer or have somebody tell or show you something that you could, drawing on current cues and your past knowledge, generate instead, you rob yourself of a powerful learning opportunity. Retrieval, in effect, is a powerful "memory modifier" (Bjork, 1975).

Closely related to the generation effect are the benefits that accompany retrieving information studied earlier. Much laboratory research (for example, Landauer & Bjork, 1978; Carrier & Pashler, 1992) has demonstrated the power of tests as learning events, and, in fact, a test or retrieval attempt, even when no corrective feedback is given, can be considerably more effective in the long term than reading material over and over. The reason why rereading is such a typical mode of studying derives, we believe, from a faulty model of how we learn and remember: We tend to think of our memories as working much like an audio/video recorder, so if we read and reread or take verbatim notes, the information will eventually write itself on our memories. Nothing, however, could be further from the way we actually learn and remember.

Unfortunately, the effectiveness of tests as learning events remains largely underappreciated, in part because testing is typically viewed as a vehicle of assessment, not a vehicle of learning. As Henry L. Roediger, Kathleen B. McDermott, and Mark A. McDaniel describe in their essay in this chapter, however, recent research using more educationally realistic materials and retention intervals has clearly demonstrated the pedagogical benefits of tests (for example, Roediger & Karpicke, 2006). Similar to the pattern with variation, spacing, and interleaving, repeated study opportunities appear, in the short term, to be more effective than repeated testing, but testing produces better recall in the long term.

Two other pedagogical benefits of tests must be mentioned: First, tests have metacognitive benefits in terms of indentifying whether information has or has not been understood and/or learned. A student's ability, for example, when going back over a chapter in a textbook, to judge whether information will be recallable on an upcoming examination is severely limited, whereas attempting to answer a fellow student's questions on the chapter can identify what has and has not been learned.

The second, related benefit is that tests can potentiate the effectiveness of subsequent study opportunities even under conditions that insure learners will be incorrect on the test (Kornell, Hays, & Bjork, 2009). Again, the basic message is that we need to spend less time restudying and more time testing ourselves.

Concluding Comments

For those of you who are students, we hope we have convinced you to take a more active role in your learning by introducing desirable difficulties into your own study activities. Above all, try to rid yourself of the idea that memory works like a tape or video recorder and that re-exposing yourself to the same material over and over again will somehow write it onto your memory. Rather, assume that learning requires an active process of interpretation—that is, mapping new things we are trying to learn onto what we already know. (There's a lesson here for those of you who are teachers—or parents—as well: Consider how you might introduce desirable difficulties into the teaching of your students or children.)

Be aware, too, when rereading a chapter or your notes, that prior exposures create a sense of familiarity that can easily be confused with understanding. And perhaps most importantly, keep in mind that retrieval—much more than restudying—acts to modify your memory by making the information you practice retrieving more likely to be recallable again in the future and in different contexts. In short, try to spend less time on the input side and more time on the output side, such as summarizing what you have read from memory or getting together with friends and asking each other questions. Any activities that involve testing yourself—that is, activities that require you to retrieve or generate information, rather than just representing information to yourself—will make your learning both more durable and flexible.

Finally, we cannot overstate the importance of learning how to manage your own learning activities. In a world that is ever more complex and rapidly changing, and in which learning on one's own is becoming ever more important, learning how to learn is the ultimate survival tool.

Suggested Further Reading

Bjork, R. A., & Bjork, E. L. (2006). Optimizing treatment and instruction: Implications of a new theory of disuse. In L-G. Nilsson and N. Ohta (Eds.), *Memory and society: Psychological perspectives* (pp. 109–133). Hove, East Sussex, England, and New York: Psychology Press.

Bjork, R. A. (1999). Assessing our own competence: Heuristics and illusions. In D. Gopher and A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 435–459). Cambridge, MA: MIT Press.

References

Bjork, R. A. (1975). Retrieval as a memory modifier. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*, pp. 123–144. Hillsdale, NJ: Erlbaum.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe and A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Eds.), *From learning processes to cognitive processes: Essays in honor of William K. Estes* (Vol. 2, pp. 35–67). Hillsdale, NJ: Erlbaum.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition, 20*, 633–642.

Kerr, R., & Booth, B. (1978). Specific and varied practice of a motor skill. *Perceptual and Motor Skills, 46*, 395–401.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science, 19*, 585–592.

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 989–998.

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory* (pp. 625–632). London: Academic Press.

Roediger, H.L., & Karpicke, J.D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255.

Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics practice problems improves learning. *Instructional Science, 35*, 481–498.

Shea, J.B., & Morgan, R.L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 179–187.

Simon, D., & Bjork, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 907–912.

Smith, S. M., Glenberg, A. M., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition, 6*, 342–353.

A Student-Centered Approach to Teaching General Biology That Really Works:

Lord's Constructivist Model Put to a TEST

PATRICIA A. BURROWES

In the comments section of a course evaluation a student once wrote, "I sleep throughout most of my Biology lectures. My professor tries hard but I just get tired of listening to so much information." This was the end of the spring semester of 1998, and I suddenly realized that I had to change the way I taught! This realization took me a long time, for in the past decade educators have been concerned with the way science is being taught, and several national efforts have been directed to redesign the instruction of pre-college science courses (AAAS, 1993, 1994; NRC, 1996). The main concern was that traditional science teaching has relied primarily on lecturing facts, and frequently requires memorization of long lists of specific vocabulary (Leonard et al., 2001). In general, the results of such teaching have been in lack of student motivation for the sciences, and limited learning reflected on poor content retention, few scientific skills, and inability to apply concepts.

PATRICIA A. BURROWES is Associate Professor in the Department of Biology at the University of Puerto Rico, San Juan, Puerto Rico 00931-3360; e-mail: pburrowe@upracd.upr.clu.edu.

When professors started to examine college instruction, they found that the same traditional teaching model was followed throughout the universities, and identified serious repercussions on the quality of science education acquired by higher level students (Adams & Slater, 1998; Anderson, 1997; Rice, 1996; Yager, 1991). Thomas Lord (1998) questioned why college students demonstrated difficulty when making connections between concepts that they had learned before, or when applying their knowledge to problem-solving situations. He thought that these problems might be a consequence of the traditional way science courses were taught, because the traditional method does not provide time for discussion, or engagement of students on inquiry-based exercises. Subsequently, Lord dramatically modified his method of instruction. Lord's innovative teaching method is student-centered and uses constructivism, active teaching, and cooperative groups. His method has proven effective in lectures and laboratories for General Biology and Environmental Science at Indiana University of Pennsylvania (Lord, 1997, 1998, 1999). Because I was impressed with the positive results of Lord's teaching techniques, and the potential application of his methods to large and small classes, I decided to test his model in my large lecture sections (100 students) of General Biology. This paper describes the results of a controlled experiment that tested the effectiveness of Lord's teaching model in:

1. Helping students achieve better grades on standard midterm exams.
2. Develop higher level thinking skills.
3. Modify their attitude towards biology at a large, urban university.

The objectives are to provide further evidence in favor of constructivist teaching over the traditional model, and to motivate fellow university professors to accept this challenge and move towards a more student-centered method of instruction.

Experimental Design

During the fall semester of 1999, I taught two large sections of General Biology I (cellular and molecular biology). One section was arbitrarily designated my control group (100 students) and was taught in the traditional manner, where instruction was based on lecturing, with little opportunity for student interaction. The other section was designated the experimental group (104 students) and taught following Lord's (1998) constructivist method (see Figure 1). Both sections met in the same large amphitheater with concentric rows of seats staggered progressively higher than the chalkboard and projection screen, which were located at the bottom and center of the room. The meeting times were equally bad for both sections of the course; the control group met Tuesdays and Thursdays from 11:30am–1:00pm (when students were starving for lunch), whereas the experimental group met Mondays and Wednesdays from 1:00pm–2:30pm (when students were sleepy from lunch). The amount of content-material covered, as well as the order it was discussed, was the same in both traditional and experimental sections.

Figure 1. Table showing the differences in teaching methods between the control and experimental sections.

TEACHING STRATEGY	
CONTROL GROUP	EXPERIMENTAL GROUP
Teacher-Centered	Student-Centered
Passive learning through teacher's lectures	Active learning through constructivist activities
No cooperative groups	Formal cooperative groups
No interaction among students	Constant interaction among students
Sporadic assessment of learning	Daily assessment of learning

Furthermore, to ensure that the amount of content was not a confounded variable in this experiment, I followed the same syllabus that had been used for the instruction of this course at our institution in previous years.

Teaching Strategies

"**Traditional teaching**" was limited to my control group where a teacher-centered environment prevailed, and course instruction emphasized content recitation, without allowing time for students to reflect upon the material presented, relate it to previous knowledge, or apply it to real life situations (Figure 1). However, natural personality attributes, such as my inborn enthusiasm for biology and tendency to modify the volume and pitch of my voice depending on the topic I am discussing, were not excluded.

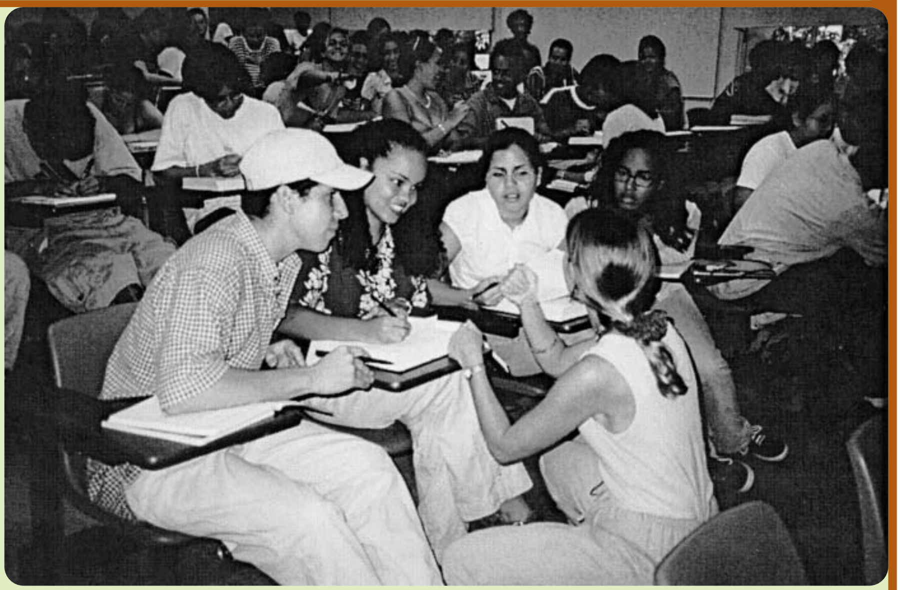
"**Experimental teaching**" was based on the Constructivist Learning Model as described by Yager (1991), the "5 E" (Engage, Explore, Explain, Elaborate, Evaluate) model developed by Bybee (1993), and cooperative learning, as modified and applied by Lord (1998, 1999, 2001). Instruction consisted of a series of short (10 to15 minutes) lectures in which I introduced new material (*Engage*), followed by the formulation of a problem or exercise (*Explore*). Depending on the task involved or the degree of difficulty, students were given 2 to 10 minutes to solve these problems with the members of their cooperative group (Figure 2). This provided an opportunity for interaction with other classmates as they tried to make sense of the new information relevant to past experiences or previous knowledge. Their consensus answers were written on a sheet that was turned in at the end of the class period (*Explain*). When the designated time to work on a problem ended, I called on 2 or 3 groups to present their answers to the rest of the class. Then, I proceeded to the *Elaborate* phase in which I either addressed misconceptions evidenced by student responses, or proceeded to tie the discussion to the introduction of new material. Listening to student responses to biological problems right after new content had been introduced provided immediate feedback on how effective my teaching was. In that sense, I had the opportunity to *Evaluate* several times within a class period. This cycle was repeated 3–5 times during each lecture period, and represented a slight variation from Lord's (1998) bookshelf approach to the "5 E" model. A final *Evaluation* occurred at the end of the class period when I received group answers to all the questions posed in class, and I offered a quiz related to the material that was discussed and worked on that day.

Organization

Pre-class planning and organization were key factors to the success of this experiment. Keeping cooperative groups throughout the semester, and being able to identify each student within a group with a color code were essential factors. On the first day of class, I used Lord's (1994) syllabus numbering to corresponding seat technique to establish formal cooperative groups in the experimental section. Groups were formed by four students seated next to each other in a row (Figure 2). After all cooperative groups were established, members were encouraged to meet one another and exchange e-mail addresses or phone numbers. Later, each group received a legal-size manila envelope that contained important information:

1. One *Cooperative Group Composition* sheet
2. Four *Student Profiles* sheets
3. One *Group Answers to Class Work* sheet
4. One *Quiz* sheet (see Figures 3–5).

Figure 2. Students working in cooperative groups during a large lecture setting of General Biology. The professor joins the group to motivate discussion.



The *Cooperative Group Composition* sheet (Figure 3) was used to make an immediate class roster organized by groups, which made grade recording much easier than using the alphabetically ordered lists provided by the Registrar office. Color assignments are related to the assessment procedure. Following Lord (1998), ten minutes prior to the termination of each class period, students were informed that it was time for a quiz. A wave of suspense took over the class. From a cloth bag I drew one of four colored balls (yellow, blue, green or red); the color drawn indicated the quiz-takers for that day. With the exception of the quiz-takers, all students (~75%) were excused from the room, leaving me with only 25 quizzes to grade after every class. The scores attained by the quiz-takers were given to all members of the cooperative group that attended class that day.

The *Student Profile* sheet (Figure 4) was completed by all students on a voluntary basis. It was designed as a data gathering instrument to test homogeneity among groups. Analysis of these data revealed that control and experimental groups were not significantly different with respect to age, sex, urban or rural origin, faculty, private or public high school education,

Figure 3. Cooperative Group Composition sheet.

COOPERATIVE GROUP COMPOSITION
GROUP #:__________________

NAME	STUDENT NUMBER	COLOR
		YELLOW
		BLUE
		RED
		GREEN

Figure 4. Student Profile sheet. Every student was provided this information at the beginning of the semester. Analysis of these data served to determine if experimental and control groups were homogeneous for comparative purposes.

STUDENT PROFILE SHEET

SECTION # ____________________

NAME: ____________________

AGE:______________ SEX: ______________

HOME TOWN: ____________________

FACULTY______________ MAJOR:______________

HIGH SCHOOL (Private or Public?)__________ HS GPA:__________

Have you taken Biology before (yes or no)?______ WHERE?__________

If answer is YES, give letter grade obtained in previous Biology course.__________

or high school grade point average (G.P.A.) and thus, were comparable. If one group had been heavily biased in any of the attributes measured (for example, 80% male in one section, versus 20% in the other), one could argue that observed differences in student academic performance among groups were due to sex composition, and not to the teaching strategy.

The forms entitled *Group Answers to Class Work* and *Quiz sheet* were distributed in the envelopes not only the first day of class, but every day thereafter (Figure 5). These forms provided a means for students to show their answers to a variety of questions and problems formulated in class, and to resolve the quiz given at the end of the period. A thoroughly worked-out *Group Answers to Class Work* sheet was worth one point for all group members who attended class that day. This was Lord's (1998) way of rewarding students for coming to class, and it worked very well. To avoid falling behind in my record keeping, I made sure that all papers were graded and corresponding scores put into a computerized roster immediately after class. This process took approximately one hour.

Two college seniors majoring in biology education contributed to the success of this project. They served as teaching assistants and attended all my classes. During the experimental section, they joined me in walking around the large amphitheater, lending help and support to cooperative groups. In addition, they participated in the design of questions and exercises for class discussion. In the control section they helped prepare, distribute, and grade quizzes, and served as unbiased critics to help me conform to "traditional" instruction methods. As the teaching assistants helped me pursue this project, they learned about the teaching strategies and action research methods.

Evaluation

In both the control and experimental sections, students could earn up to 400 points (perfect 100% grade) throughout the semester. However, the distribution of points for in-class work, quizzes, exams, and laboratory work varied slightly among control and experimental groups (Figure 6). In-class assessment for the control (i.e., traditional) group was based on eight 10-point pop-quizzes which all students were required to take. Students who were always present when a quiz was given were more likely to attain the 60 maximum quiz points considered for grade computation, and could earn up to 5 more points as a bonus. In this manner, control students were encouraged to attend class every day. When students in the experimental group attended class for more than 20 periods, they were rewarded by the opportunity to accumulate up to 5 extra points via daily *Group Answers to Class Work* sheets. I wrote my own midterm exams (not those of a departmental committee), in multiple choice fashion. Exams consisted of approximately 50% content recall and 50% conceptual understanding/application questions. For comparative purposes, the same tests were administered to both experimental and control groups on the same nights during the semester. Additionally, all students took the same comprehensive, departmental final exam during the last week of the semester. This was a concession to

continued on page 496

the department, to assure other faculty members that I was going to cover all the content material usually taught in this course.

Academic Achievement Results

Exams were taken on computer sheets and graded electronically. Grades achieved by students in experimental (constructivist teaching) and control (traditional teaching) groups were contrasted graphically (Figure 7) and the mean test scores were compared statistically by students' T-tests using Minitab 12 software. The first partial exam was offered after six weeks of instruction and included content on the cell as the functional unit of life (atoms, molecules, the cell membrane, organelles, energy transformations, cellular respiration, and photosynthesis). Although average scores of students in the experimental group were significantly better than in the control group ($\bar{x}$ = 65% *versus* 58%; T = 2.65, P = 0.004, n = 204), and they attained more As and Bs, and fewer Fs, the differences are not as impressive as later in the semester (Figure 7). The second exam was given 12 weeks into the semester, and evaluated knowledge on the continuity of life (mitosis, meiosis, DNA structure and replication, protein synthesis, and inheritance). Results of this exam showed grade improvement in both groups (Figure 7); however, the mean score of students in the experimental group was significantly higher than that of students in the control section ($\bar{x}$ = 72% *versus* 67%; T = 2.41, P = 0.009, n = 192). The outcome of the third exam (evolution and origin of life) was striking because performance of students in the experimental group approximated an ideal normal distribution of grades (Figure 7). Although students' achievement in the control group improved, students in the experimental section still did significantly better ($\bar{x}$ = 74% *versus* 68%; T = 3.05, P = 0.001 n =190). For the final exam, it is common policy at our institution to excuse students who by the end of the semester have attained an "A" average in the course. As a consequence, when comparing final exam grades between experimental and control groups (Figure 8), the best students were not considered. Mean final exam grades were not significantly different between sections ($\bar{x}$ = 67% *versus* 64%; T = 1.29, P = 0.20, n =172); however, more students earned As and Bs in the experimental group than in the control. Thus, it is clear that students instructed in a constructivist, active-learning

Figure 5. Evidence of Group Effort in Class Work and Quiz sheets were distributed daily in legal-sized envelopes to each cooperative group of the experimental section. These were blank sheets with the following headings:

GROUP ANSWERS TO CLASS WORK

Group # ______________________ Date ______________________

Members Present:

______________________ ______________________

______________________ ______________________

QUIZ SHEET

Group # ______________________ Date ______________________

Quiz Taker: Color:

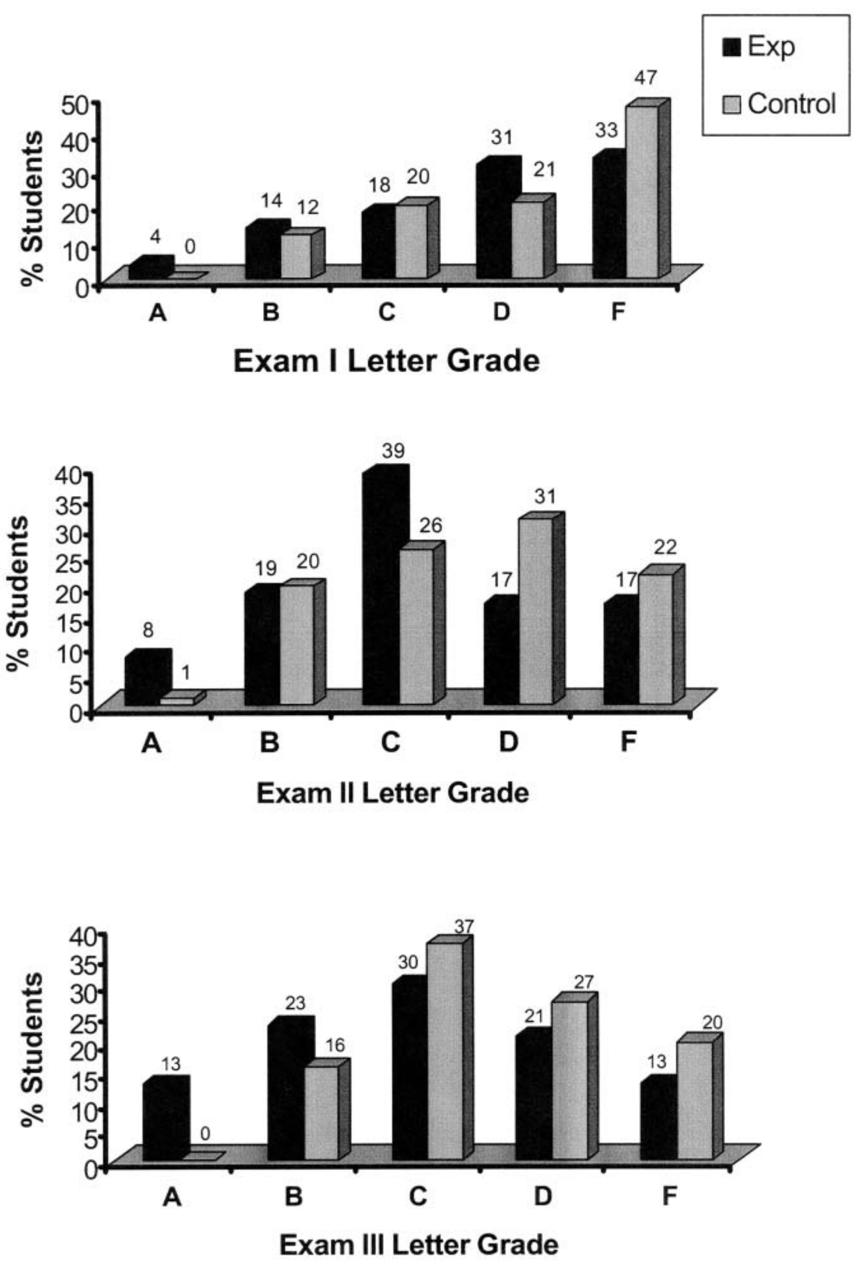
______________________ ______________________

Figure 6. Table showing the differences on the evaluation criteria for control and experimental groups.

EVALUATION

CONTROL GROUP		EXPERIMENTAL GROUP	
3 Midterm Exams	180 pts.	3 Midterm Exams	180 pts.
Final Exam	60 pts.	Final Exam	60 pts.
8 Pop Quizzes - 10 pts. each	60 pts.	Quizzes - 2 pts. each, 1 per class	40 pts.
		Group Class Work	20 pts.
Laboratory	100 pts.	Laboratory	100 pts.
TOTAL	400 pts.	TOTAL	400 pts.

Figure 7. Comparative and progressive letter grades achieved by experimental (constructivist teaching) and control (traditional teaching) sections in the three midterm exams offered during a General Biology I course.

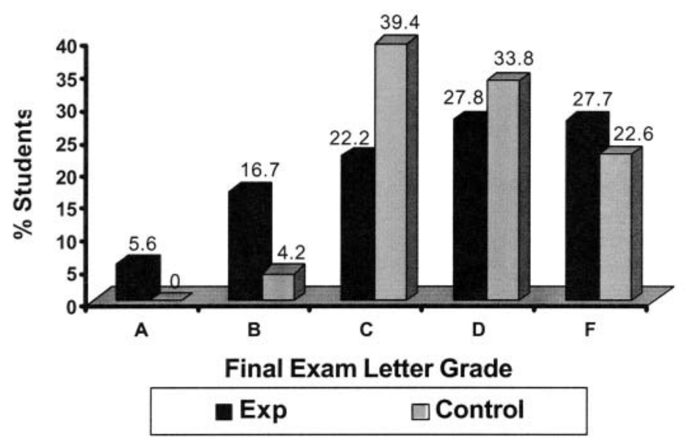


environment were able to perform better on the same tests than students taught in a traditional fashion. The drop in grades in the comprehensive final exam is a typical phenomenon that I have observed for several years. It may be related to pressure from other finals, lack of ability to deal with large content material, attention drifting to vacation activities, or simply exhaustion.

Conceptual Understanding Results

To train students to become scientists, we must provide opportunities to participate in all aspects of the scientific method. By participating in the scientific process, students learn to think scientifically (Marbach-Ad & Claassen, 2001). One of the benefits of Lord's (1998) constructivist model is that it offers students many chances to develop higher-order thinking skills through a variety of in-class exercises – the idea is that practice makes excellence. I refer to higher order thinking skills here as the ability to think critically, make reasoned judgements about complex issues or data, and apply concepts to resolve problems or investigate questions. Examples of exercises that help develop these skills are multi-answer questions, concept maps, discussion scenarios, graph interpretation, graphing of data, drawing conclusions, solving problems, etc. Concrete examples of some of these activities for biology are listed in Appendix I, but more can be found in Lord (1998) and Burrowes and Nazario (2001). Other curriculum models have been developed to improve thinking skills among college students (Crow, 1989; Lawson, 1992, 1999; Pheeney, 1997), and recent literature provides examples of how to reform laboratories to an inquiry-based format allowing students to learn science by inquiry (Marbach-Ad & Claassen, 2001; Sundberg et al., 2000). Having provided such an environment in the experimental section

Figure 8. Comparative results of grades attained by experimental and control groups in the comprehensive final exam offered at the end of a General Biology I course.



Results from the pre-evaluation helped identify problems before it was too late to solve them. Some students in the experimental group commented that not all group members participated actively or did equally well on the quizzes, so I took the time to give them a brief "pep-talk" on real cooperative work. Others mentioned that I spoke too fast, and still others complained about not having enough time to solve some of the problems given in class. All these issues were attended to immediately. Another asset of the pre-evaluation was that it revealed that students from both groups rated highly my performance as a teacher. This implied that students in the experimental section liked the change, and that control-section students felt comfortable with my "traditional" instruction. Once I had the results of the post-evaluation, I compared students' response to the question, "How would you rate your interest in biology — high, medium or low?" At the beginning of the semester, the majority of the students answered "medium" to this question, and responses were independent from experimental vs. control sections. At the end of the semester, there were differences in attitude toward biology that could be significantly associated with groups ($X^2 = 6.52$, $P = 0.04$, $df = 2$); more students in the experimental group (70% vs. 50%) expressed that their interest in biology was high. This kind of change in perception suggests that a student-centered, active learning environment was more effective at motivating students to become interested in biology than a teacher-centered, passive approach.

only (Figure 1), I hypothesized that these students should do better than control-group students on exam questions that addressed higher-order thinking skills. To test this hypothesis I selected 21 questions (seven from each of the three midterm exams) that required interpretation of graphs or data, application of concepts to solve problems, establishing connections among related topics, or the ability to make inferences from given facts. When the answers to these questions were compared among experimental and control group students, a significantly greater number of experimental group students provided the correct answer ($T = 3.79$, $P = 0.001$, $n = 21$), suggesting that, indeed, in-class practice of problem-solving techniques does help to develop skills for scientific thinking.

Attitude Assessment

An identical course evaluation that included questions related to students' interest towards biology, impression of the classroom setting, and the professor's work was given to control and experimental sections after the first midterm exam, and again during the last day of class. The objective was threefold:

1. To obtain preliminary information on how students felt about the class and my teaching.
2. To determine if I was being fair to both groups.
3. To record if and how students' opinions changed by the end of the course.

There were other interactions that took place during the semester that revealed a more assertive, intellectual, and competent attitude when discussing science among students from the experimental group. For example, the day after every midterm exam, I brought copies of the test to class and distributed one copy to every cooperative group in the experimental section, and one copy to every four students that happened to sit together that day in the control section. Working in groups, students had the opportunity to solve the test for a second time, but with only half the time provided during the original exam. If they were able to obtain a perfect score, each student from that group would receive 3 bonus points (to be added to the score obtained on the test); those who scored 95-99% could add 2 points to their original test, and those who scored 90-94% could add 1 extra point. More groups obtained bonus points in the experimental section than in the control section of each of the three exams (Exam I: 5 vs. 1; Exam II: 8 vs. 3; and Exam III: 10 vs.

6). In addition, the discussion that took place within groups was more lively and led students to challenge the way questions were written or the possibility of alternate answers much more frequently in the experimental section than in the control. This attitude difference can be explained by the fact that in the constructivist-active learning environment, the groups were established at the beginning of the semester and students had been given many opportunities to discuss and interact in a cooperative fashion. On the other hand, students taught in a teacher-centered traditional manner did not necessarily know their neighboring classmates and were not accustomed to discussing biology in a group (Figure 1).

Frequently Asked Questions

I have had the chance to listen to Tom Lord present talks on his constructivist teaching strategy to biology scholars on three occasions, and I have twice presented a seminar on my own work to colleagues. As I have heard people ask the same questions every time, I decided to include a "Frequently Asked Questions" (FAQ) section. Hopefully, I will be able to address some of the general concerns about problems that may arise when using this technique, and how to go about solving them.

1. *How do you manage the time allocated to problem solving in class so that it does not get out of hand?*

   I use a stop watch with a loud bell (such as those used to keep baking times in the kitchen). Once I have told the students how much time they have for a particular exercise, I set the stop watch and when the bell rings, I immediately move on to the discussion.

2. *Did you fall behind on content coverage in your experimental section with respect to the control group?*

   No, not even for the first exam when I was still inexperienced. This teaching approach does not require more class time, it just cuts from the time the teacher is speaking, and gives it to the students to question, discover, and learn on their own.

3. *What happens if the student who is supposed to take the quiz (according to the colored-ball that was drawn) is absent that day?*

   I draw another ball, and another if necessary, until all groups have a quiz-taker present.

4. *What do you do with a group that is upset because it has a particular quiz-taker who always flunks the quiz?*

First, I remind the unhappy students that they have a legitimate problem and that they should try to solve it by talking to the culprit themselves. If that does not work, I will intercede. This problem has never gone beyond a private chat between me and a student about responsibility and obligation to his/her group.

5. *How did the number of students dropping out of the course compare between experimental and control groups?*

   Four experimental-group students dropped the course after the first exam, whereas 12 students dropped out of the control section at different times during the semester.

6. *How did the number of students excluded from the final exam for having an A average in the course compare between experimental and control groups?*

   As expected from the partial exam results, more students from the experimental group (12) were excused from the final exam than from the control group (4).

Conclusions & Implications

This study provides substantiated evidence that teaching in a constructivist, active learning environment is more effective than traditional instruction in promoting academic achievement, increasing conceptual understanding, developing higher level thinking skills, and enhancing students interest in biology. In their final course evaluations, students in the experimental section commented that they enjoyed this class much more than their traditional classes, felt they had learned more, made valuable friendships in their collaborative groups and – particularly important to me – they never fell asleep! Thus, I am convinced that constructivism works better for our generation of students, and I will never return to a traditional style of teaching. Although the constructivist method of instruction requires a greater investment of time and effort from the professor for preparation, organization, and grading, the majority of this investment is made the first semester of teaching. During subsequent semesters, effort/payback increases dramatically, as less time is required. For example, with experience, I have become more efficient at formulating questions and coming up with ideas for problems, scenarios, and case studies, which help students develop their own knowledge of the material. Additionally, help from trained teaching assistants in grading, book-keeping, and organizational tasks associated with instruction can reduce some of the workload required of the instructor. Suggestions to help professors and students make the transition from a teacher-centered to a student-centered learning environment, where pupils assume responsibility for their education, are available in the literature (Lord, 1998; Modell, 1996). In most cases, I recommend this change to be an on-going process in which professors experiment and evaluate techniques to find out which make them most comfortable. As a result, educators will engage in active research – an excellent tool to learn about the effectiveness of our teaching and how our particular population of students learn best. I have been applying this constructivist, student-centered teaching method to my Biology and Zoology classes for the past four semesters, and I am happier as a professor, not only because I enjoy my teaching experience much more, but because the overall results of empowerment among my students is overwhelming.

Acknowledgments

This study would not have been possible without the collaboration of all the students enrolled in my Biology 3101 sections the fall semester of 1999, who voluntarily agreed to participate as research subjects. Miley Oliveras and Manuela Ortíz dedicated more than their share of time as research assistants in this project. I thank Tom Lord for providing inspiration and useful comments to improve earlier versions of this manuscript, and Gladys M. Nazario for her pedagogic advice and constant support of my work. This research was funded in part by a grant from the Collaborative for Research in Teacher Preparation (PR-CETP) to G. M. Nazario and myself, and by the Department of Biology at the University of Puerto Rico.

References

Adams, J. P. & Slater, T. F. (1998). Using action research to bring the large class down to size. *Journal of College Science Teaching,* November 1998.

American Association for the Advancement of Science (AAAS). (1993). *Benchmarks for Science Literacy.* New York: Oxford University Press.

American Association for the Advancement of Science (AAAS). (1994). *Science for all Americans.* New York: Oxford University Press.

Anderson, E. (1997). Active learning in the lecture hall. *Journal of College Science Teaching, 26*(6), 428–429.

Burrowes, P. A. & Nazario, G. M. (2001). Preparing students for the transition from a teacher-centered to a student-centered environment: Active exercises that work at the university level. *Pedagogía, 35,* 135-141.

Bybee, R. (1993). *Instructional Model for Science Education, in Developing Biological Literacy.* Colorado Springs, CO: Biological Sciences Curriculum Studies.

Crow, L. W. (1989). The nature of critical thinking. *Journal of College Science Teaching, 19,* 114-116.

Lawson, A. E. (1992). The development of reasoning among college biology students — a review of research. *Journal of College Science Teaching, 21,* 338–344.

Lawson, A. E. (1999). What should students learn about the nature of science and how should we teach it? *Journal of College Science Teaching, 28,* 401–411.

Leonard, W. H., Speziale, B. J. & Penick, J. (2001). Performance assessment of a standards-based high school biology curriculum. *The American Biology Teacher, 63*(5), 310–316.

Lord, T. (1994). Using constructivism to enhance student learning in college biology. *Journal of College Science Teaching, 5,* 346–348.

Lord, T. (1997). Comparing traditional and constructivist teaching in college biology. *Innovative Higher Education, 21*(3), 197–217.

Lord, T. (1998). Cooperative learning that really works in biology teaching. Using constructivist-based activities to challenge student teams. *The American Biology Teacher, 60*(8), 580–588.

Lord, T. (1999). A comparison between traditional and constructivist teaching in environmental science. *The Journal of Environmental Education, 30*(3), 22–28.

Lord, T. (2001). 101 reasons for using cooperative learning in biology teaching. *The American Biology Teacher, 63*(1), 30–38.

Marbach-Ad, G. & Claassen, L.A. (2001). Improving students' questions in inquiry labs. *The American Biology Teacher, 63*(6), 410–419.

Minitab. (1998). Statistical Software for Windows. Release 12.

Modell, H. I. (1996). Preparing students to participate in an active learning environment. *Advances in Physiology Education, 15*(1), S69–S77.

National Research Council (NRC). (1996). *National Science Education Standards.* Washington, DC: National Academy Press.

Pheeney, P. (1997). Hands on, minds on. Activities to engage our students. *Science Scope,* September 1997, 30–33.

Rice, R. E. (1996). *Making a Place for the New American Scholar.* Washington DC: American Association for Higher Education.

Sundberg, M.D., Armstrong, J.E., Dini, M.L. & Wischusen, E.W. (2000). Some practical ways for instituting investigative biology laboratories. *Journal of College Science Teaching,* March/April 2000, 353–359.

Yager, R. (1991). The constructivist learning model: Toward real reform in science education. *The Science Teacher,* September, 53–57.

Appendix I. Examples of exercises that can be given to student groups to work in class. Approximate time to be allocated is in parenthesis.

Concept maps. After discussing evolution, ask students to draw a concept map to connect the following terms: *evolution, gene bank, variability, natural selection, mutations, recombination, gene flow, selection pressures, genetic drift,* and *reproductive potential.* Every term should be connected with arrows labeled with a word that describes the link between the processes – example: change, causes, provides, directs, etc. (15 minutes)

Scenarios. When you are discussing homeostasis, tell the students that you just received a VIP letter asking for help from your students in solving a problem of unknown purpose that concerns "National Security." The students are now part of a multi-disciplinary team put together to design "the perfect animal" that can survive and reproduce successfully under the following conditions: an environment that is very hot and dry during the day, but turns cold and windy at night, and that has many fast and aggressive predators. In their design of this animal, they should consider integument, body support, reproductive strategy, excretion and mode of locomotion. (15 minutes)

Graphs. After studying animal circulation, ask students to draw a graph showing the relationship between blood pressure (mmHg) or velocity of flow, and the diameter of blood vessels. Depending on your group of students, you may want to provide some additional instructions like how to set the axis, and then give them the freedom to do a projection, curve or bar diagram. (5 minutes)

Brainstorming. When reviewing thermoregulation, ask students to come up with five ways in which snakes can prevent overheating on a hot summer day. (3 minutes)

Observations/Predictions. When you are about to explain the cell membrane models, show an illustration of the fluid mosaic model and ask students to make a list of six observations. If they do not recognize a particular structure, they can describe it using terms like "blob," "cylinder," or "ball." (3 minutes) When you discuss the fluid mosaic model with the students, they have already spent time studying its structure and will be more receptive to learn the proper names of the molecules and their function. Then, ask students to predict what would happen to the permeability of the cell membrane if the proteins were removed. (5 minutes)

Problems. This problem was designed to help students understand the difference between passive and active transport. A famous musician (make it relevant – I used Ricky Martin) is giving a concert in town. People are camping out to purchase tickets the night before the ticket counter is to open. There is a single, very long line to purchase tickets until, suddenly, a second window opens for sales. Predict:

1. What will happen?
2. What will happen if one of the two lines gets a little longer than the other?
3. What will you need to do to get people to move from a shorter to a longer line? (5 minutes)

Data to graph. On enzyme action: Mothers complained that their children would not eat the fresh apple wedges they sent to school in lunch boxes because they turned "brown and disgusting." Discuss with your students what causes the change of color in apples when they are exposed to air. The following data were collected during an experiment. Graph the data, interpret the results of the experiment, and make concluding remarks regarding the best way to pack fresh apple wedges in students' lunch boxes. The results of graphing these data will set the basis for a follow-up discussion on enzyme optimal and saturation levels. (15 minutes)

	Color of apple from normal =1, to very brown and disgusting =5		
Time (minutes)	Wedges *alone*	Wedges *in lime juice*	Wedges *in ice*
0	1	1	1
5	2	1	1
10	3	1	1
15	4	1	2
20	5	1	2
25	5	1	2
30	5	1	3

Acknowledgments. This work was supported by the Belgian Fonds National de la Recherche Scientifique (FNRS), European Commission, Mind Science Foundation, McDonnell Foundation, French-Speaking Community Concerted Research Action (ARC 06/11-340), Fondation Léon Frédéricq, and National Institutes of Health. M.-A.B. and O.G. are Research Fellows, M.B. and C.S. Postdoctoral Fellows, and S.L. Senior Research Associate at the FNRS. M.I.G., V.L., and K.F. are supported by the Wellcome Trust.

Supporting Online Material
www.sciencemag.org/cgi/content/full/332/6031/858/DC1
Materials and Methods
SOM Text
Fig. S1
Table S1
References

22 December 2010; accepted 6 April 2011
10.1126/science.1202043

Improved Learning in a Large-Enrollment Physics Class

Louis Deslauriers,[1,2] Ellen Schelew,[2] Carl Wieman*†‡

We compared the amounts of learning achieved using two different instructional approaches under controlled conditions. We measured the learning of a specific set of topics and objectives when taught by 3 hours of traditional lecture given by an experienced highly rated instructor and 3 hours of instruction given by a trained but inexperienced instructor using instruction based on research in cognitive psychology and physics education. The comparison was made between two large sections (N = 267 and N = 271) of an introductory undergraduate physics course. We found increased student attendance, higher engagement, and more than twice the learning in the section taught using research-based instruction.

The traditional lecture approach remains the prevailing method for teaching science at the postsecondary level, although there are a growing number of studies indicating that other instructional approaches are more effective (*1*–*8*). A typical study in the domain of physics demonstrates how student learning is improved from one year to the next when an instructor changes his or her approach, as measured by standard concept-based tests such as the Force Concept Inventory (*9*) or the instructor's own exams. In our studies of two full sessions of an advanced quantum mechanics class taught either by traditional or by interactive learning style, students in the interactive section showed improved learning, but both sections, interactive and traditional, showed similar retention of learning 6 to 18 months later (*10*). Here, we compare learning produced by two contrasting instructional methods in a large-enrollment science course. The control group was lectured by a motivated faculty member with high student evaluations and many years of experience teaching this course. The experimental group was taught by a postdoctoral fellow using instruction based on research on learning. The same selected learning objectives were covered by both instructors in a 1-week period.

The instructional design for the experimental section was based on the concept of "deliberate practice" (*11*) for the development of expertise. The deliberate practice concept encompasses the educational ideas of constructivism and formative assessment. In our case, the deliberate practice takes the form of a series of challenging questions and tasks that require the students to practice physicist-like reasoning and problem solving during class time while provided with frequent feedback.

The design goal was to have the students spend all their time in class engaged in deliberate practice at "thinking scientifically" in the form of making and testing predictions and arguments about the relevant topics, solving problems, and critiquing their own reasoning and that of others. All of the activities are designed to fit together to support this goal, including moving the simple transfer of factual knowledge outside of class as much as possible and creating tasks and feedback that motivate students to become fully engaged. As the students work through these tasks, they receive feedback from fellow students (*12*) and from the instructor. We incorporate multiple "best instructional practices," but we believe the educational benefit does not come primarily from any particular practice but rather from the integration into the overall deliberate practice framework.

This study was carried out in the second term of the first-year physics sequence taken by all undergraduate engineering students at the University of British Columbia. This calculus-based course covers various standard topics in electricity and magnetism. The course enrollment was 850 students, who were divided among three sections. Each section had 3 hours of lecture per week. The lectures were held in a large theater-style lecture hall with fixed chairs behind benches grouping up to five students. The students also had weekly homework assignments, instructional laboratories, and tutorials and recitations where they solved problems; this work was graded. There were two midterm exams and a final exam. All course components were common across all three sections, except for the lectures, which were prepared and given independently by three different instructors.

During week 12, we studied two sections whose instructors agreed to participate. For the 11 weeks preceding the study, both sections were taught in a similar manner by two instructors (A and B), both with above average student teaching evaluations and many years experience teaching this course and many others. Both instructors lectured using PowerPoint slides to present content and example problems and also showed demonstrations. Meanwhile, the students took notes. "Clicker" (or "personal response system") questions (average 1.5 per class, range 0 to 5) were used for summative evaluation (which was characterized by individual testing without discussion or follow-up other than a summary of the correct answers). Students were given participation credit for submitting answers.

Before the experiment, a variety of data were collected on the students in the two sections

Table 1. Measures of student perceptions, behaviors, and knowledge.

	Control section	Experimental section
Number of students enrolled	267	271
Mean BEMA score (*13*) (week 11)	47 ± 1%	47 ± 1%
Mean CLASS score (*14*) (start of term) (agreement with physicist)	63 ± 1%	65 ± 1%
Mean midterm 1 score	59 ± 1%	59 ± 1%
Mean midterm 2 score	51 ± 1%	53 ± 1%
Attendance before experiment*	55 ± 3%	57 ± 2%
Attendance during experiment	53 ± 3%	75 ± 5%
Engagement before experiment*	45 ± 5%	45 ± 5%
Engagement during experiment	45 ± 5%	85 ± 5%

*Average value of multiple measurements carried out in a 2-week interval before the experiment. Engagement also varies over location in the classroom; numbers given are spatial and temporal averages.

[1]Carl Wieman Science Education Initiative, University of British Columbia, Vancouver, BC, Canada. [2]Department of Physics and Astronomy, University of British Columbia, Vancouver, BC, Canada.

*On leave from the University of British Columbia and the University of Colorado.

†To whom correspondence should be addressed. E-mail: gilbertwieman@gmail.com

‡This work does not necessarily represent the views of the Office of Science and Technology Policy or the United States government.

(Table 1). Students took two midterm exams (identical across all sections). In week 11, students took the Brief Electricity and Magnetism Assessment (BEMA), which measures conceptual knowledge (*13*). At the start of the term, students took the Colorado Learning Attitudes about Science Survey (CLASS) (*14*), which measures a student's perceptions of physics. During weeks 10 and 11, we measured student attendance and engagement in both sections. Attendance was measured by counting the number of students present, and engagement was measured by four trained observers in each class using the protocol discussed in the supporting online material (SOM) (*15*). The results show that the two sections were indistinguishable (Table 1). This in itself is interesting, because the personalities of the two instructors are rather different, with instructor A (control section) being more animated and intense.

The experimental intervention took place during the 3 hours of lecture in the 12th week. Those classes covered the unit on electromagnetic waves. This unit included standard topics such as plane waves and energy of electromagnetic waves and photons. The control section was taught by instructor A using the same instructional approach as in the previous weeks, except they added instructions to read the relevant chapter in the textbook before class. The experimental section was taught by two instructors who had not previously taught these students. The instructors were the first author of this paper, L.D., assisted by the second author, E.S. Instructor A and L.D. had agreed to make this a learning competition. L.D. and instructor A agreed beforehand what topics and learning objectives would be covered. A multiple-choice test (see SOM) was developed by L.D. and instructor A that they and instructor B agreed was a good measure of the learning objectives and physics content. The test was prepared at the end of week 12. Most of the test questions were clicker questions previously used at another university, often slightly modified. Both sections were told that they would receive a bonus of 3% of the course grade for the combination of participating in clicker questions, taking the test, and (only in the experimental section) turning in group task solutions, with the apportionment of credit across these tasks left unspecified.

In contrast to instructor A, the teaching experience of L.D. and E.S. had been limited to serving as teaching assistants. L.D. was a postdoctoral researcher working in the Carl Wieman (third author of this paper) Science Education Initiative (CWSEI) and had received training in physics education and learning research and methods of effective pedagogy while assisting with the teaching of six courses. E.S. had a typical physics graduate student background except for having taken a seminar course in physics education.

The instructional approach used in the experimental section included elements promoted by CWSEI and its partner initiative at the University of Colorado: preclass reading assignments, preclass reading quizzes, in-class clicker questions with student-student discussion (CQ), small-group active learning tasks (GT), and targeted in-class instructor feedback (IF). Before each of the three 50-min classes, students were assigned a three- or four-page reading, and they completed a short true-false online quiz on the reading. To avoid student resistance, at the beginning of the first class, several minutes were used to explain to students why the material was being taught this way and how research showed that this approach would increase their learning.

A typical schedule for a class was the following: CQ1, 2 min; IF, 4 min; CQ2, 2 min; IF, 4 min; CQ2 (continued), 3 min; IF, 5 min; Revote CQ2, 1 min; CQ3, 3 min; IF, 6 min; GT1, 6 min; IF with a demonstration, 6 min; GT1 (continued), 4 min; and IF, 3 min. The time duration for a question or activity includes the amount of time the students spent discussing the problem and asking numerous questions. There was no formal lecturing; however, guidance and explanations were provided by the instructor throughout the class. The instructor responded to student-generated questions, to results from the clicker responses, and to what the instructor heard by listening in on the student-student discussions. Students' questions commonly expanded upon and extended the material covered by the clicker questions or small-group tasks. The material shown on the slides used in class is given in the SOM, along with some commentary about the design elements and preparation time required.

At the beginning of each class, the students were asked to form groups of two. After a clicker question was shown to the class, the students discussed the question within their groups (which often expanded to three or more students) and submitted their answer using clickers. When the voting was complete, the instructor showed the results and gave feedback. The small-group tasks were questions that required a written response. Students worked in the same groups but submitted individual answers at the end of each class for participation credit. Instructor A observed each of these classes before teaching his own class and chose to use most of the clicker questions developed for the experimental class. However, Instructor A used these only for summative evaluation, as described above.

L.D. and E.S. together designed the clicker questions and small-group tasks. L.D. and E.S. had not taught this class before and were not familiar with the students. Before the first class, they solicited two volunteers enrolled in the course to pilot-test the materials. The volunteers were asked to think aloud as they reasoned through the planned questions and tasks. Results from this testing were used to modify the clicker questions and tasks to reduce misinterpretations and adjust the level of difficulty. This process was repeated before the second class with one volunteer.

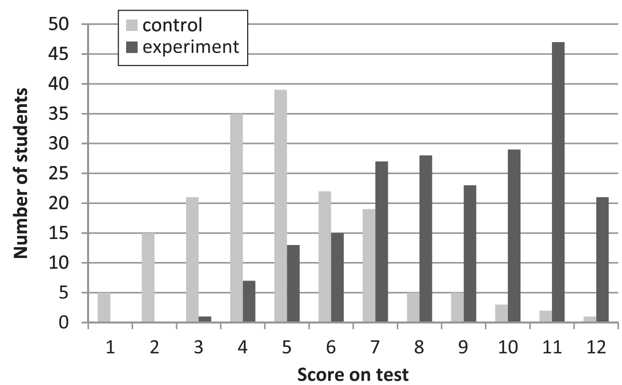
During the week of the experiment, engagement and attendance remained unchanged in the control section. In the experimental section, student engagement nearly doubled and attendance increased by 20% (Table 1). The reason for the attendance increase is not known. We hypothesize that of the many students who attended only part of a normal class, more of them were captured by the happenings in the experimental section and decided to stay and to return for the subsequent classes.

The test was administered in both sections in the first class after the completion of the 3-hour unit. The control section had covered the material related to all 12 of the questions on the test. The experimental section covered only 11 of the 12 questions in the allotted time. Two days before the test was given, the students in both sections were reminded of the test and given links to the postings of all the material used in the experimental section: the preclass reading assignments and quizzes; the clicker questions; and the group tasks, along with answers to all of these. The students were encouraged by e-mail and in class to try their best on the test and were told that it would be good practice for the final exam, but their performance on the test did not affect their course grade. Few students in either section finished in less than 15 min, with the average being about 20 min.

The test results are shown in Fig. 1. For the experimental section, 211 students attended class to take the test, whereas 171 did so in the control section. The average scores were 41 ± 1% in the control section and 74 ± 1% in the experimental section. Random guessing would produce a score of 23%, so the students in the experimental section did more than twice as well on this test as those in the control section.

The test score distributions are not normal (Fig. 1). A ceiling effect is apparent in the experi-

Fig. 1. Histogram of student scores for the two sections.



mental section. The two distributions have little overlap, demonstrating that the differences in learning between the two sections exist for essentially the entire student population. The standard deviation calculated for both sections was about 13%, giving an effect size for the difference between the two sections of 2.5 standard deviations. As reviewed in (*4*), other science and engineering classroom studies report effect sizes less than 1.0. An effect size of 2, obtained with trained personal tutors, is claimed to be the largest observed for any educational intervention (*16*).

This work may obtain larger effect sizes than in this previous work because of the design and implementation that maximized productive engagement. The clicker questions and group tasks were designed not only to require explicit expert reasoning but also to be sufficiently interesting and personally relevant to motivate students to fully engage. Another factor could be that previous work primarily used end-of-term tests, and the results on those tests reflect all the learning that students do inside and outside of class, for example, the learning that takes place while doing homework and studying for exams. In our intervention, the immediate low-stakes test more directly measured the learning achieved from preclass reading and class itself, in the absence of subsequent study.

We are often asked about the possible contributions of the Hawthorne effect, where any change in conditions is said to result in improved performance. As discussed in citations in the SOM, the original Hawthorne plant data actually show no such effect, nor do experiments in educational settings (*17*).

A concern frequently voiced by faculty as they consider adopting active learning approaches is that students might oppose the change (*18*). A week after the completion of the experiment and exam, we gave students in the experimental section an online survey (see SOM); 150 students completed the survey.

For the survey statement "I really enjoyed the interactive teaching technique during the three lectures on E&M waves," 90% of the respondents agreed (47% strongly agreed, 43% agreed) and only 1% disagreed. For the statement "I feel I would have learned more if the whole physics 153 course would have been taught in this highly interactive style." 77% agreed and only 7% disagreed. Thus, this form of instruction was well received by students.

In conclusion, we show that use of deliberate practice teaching strategies can improve both learning and engagement in a large introductory physics course as compared with what was obtained with the lecture method. Our study compares similar students, and teachers with the same learning objectives and the same instructional time and tests. This result is likely to generalize to a variety of postsecondary courses.

References and Notes

1. R. J. Beichner *et al.*, in *Research-Based Reform of University Physics*, E. F. Redish, P. J. Cooney, Eds. (American Association of Physics Teachers, College Park, MD, 2007).
2. C. H. Crouch, E. Mazur, *Am. J. Phys.* **69**, 970 (2001).
3. J. E. Froyd, "White paper on promising practices in undergraduate STEM education" [Commissioned paper for the Evidence on Promising Practices in Undergraduate Science, Technology, Engineering, and Mathematics (STEM) Education Project, The National Academies Board on Science Education, 2008]. www7.nationalacademies.org/bose/Froyd_Promising_Practices_CommissionedPaper.pdf
4. J. E. Froyd, "Evidence for the efficacy of student-active learning pedagogies" (Project Kaleidoscope, 2007). www.pkal.org/documents/BibliographyofSALPedagogies.cfm
5. R. R. Hake, *Am. J. Phys.* **66**, 64 (1998).
6. J. K. Knight, W. B. Wood, *Cell Biol. Educ.* **4**, 298 (2005).
7. M. Prince, *J. Eng. Educ.* **93**, 223 (2004).
8. L. Springer, M. E. Stanne, S. S. Donavan, *Rev. Educ. Res.* **69**, 21 (1999).
9. D. Hestenes, M. Wells, G. Swackhamer, *Phys. Teach.* **30**, 141 (1992).
10. L. Deslauriers, C. Wieman, *Phys. Rev. ST Phys. Educ. Res.* **7**, 010101 (2011).
11. K. A. Ericsson, R. Krampe, C. Tesch-Romer, *Psychol. Rev.* **100**, 363 (1993).
12. M. K. Smith *et al.*, *Science* **323**, 122 (2009).
13. L. Ding, R. Chabay, B. Sherwood, R. Beichner, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010105 (2006).
14. W. K. Adams *et al.*, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010101 (2006).
15. Materials and methods are available as supporting material on *Science* Online.
16. B. Bloom, *Educ. Res.* **13**, 4 (1984).
17. R. H. Bauernfeind, C. J. Olson, *Phi Delta Kappan* **55**, 271 (1973).
18. G. K. Allen, J. F. Wedman, L. C. Folk, *Innovative High. Educ.* **26**, 103 (2001).

Acknowledgments: This work was supported by the University of British Columbia through the Carl Wieman Science Education Initiative.

Supporting Online Material

www.sciencemag.org/cgi/content/full/332/6031/862/DC1
Materials and Methods
SOM Text
References

16 December 2010; accepted 5 April 2011
10.1126/science.1201783

www.sciencemag.org/cgi/content/full/332/6031/862/DC1

Supporting Online Material for

Improved Learning in a Large-Enrollment Physics Class

Louis Deslauriers, Ellen Schelew, Carl Wieman*

*To whom correspondence should be addressed. E-mail: gilbertwieman@gmail.com

Published 13 May 2011, *Science* **332**, 862 (2011)
DOI: 10.1126/science.1201783

This PDF file includes:

Materials and Methods
SOM Text
References

Supporting Online Material for

Improved learning in a large enrollment physics class

L. Deslauriers, E. Schelew, C. Wieman

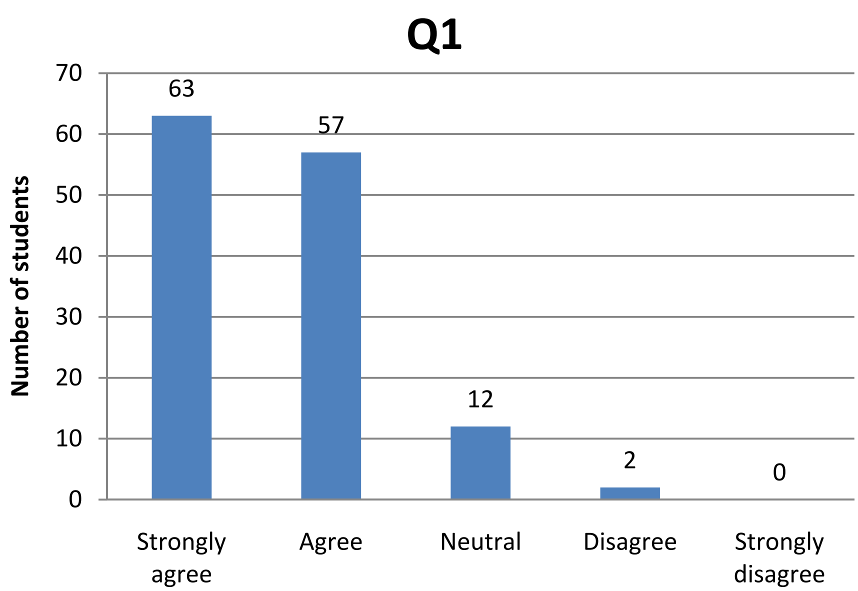
1. Engagement measurements discussion.
2. Experimental section opinion survey and responses.
3. Test given to both sections on the material taught.
4. Slides shown in the three days of class in the experimental section. There is typically one question or task per slide, with about six slides per 50 minute class. Commentary on the design and preparation is inserted (in italics).
5. Learning objectives agreed upon by the two instructors.
6. Hawthorne effect comment, and discussion of engagement and attendance in courses with similar design over a full semester.
7. List of proven teaching practices used, with references.

1. Engagement measurements

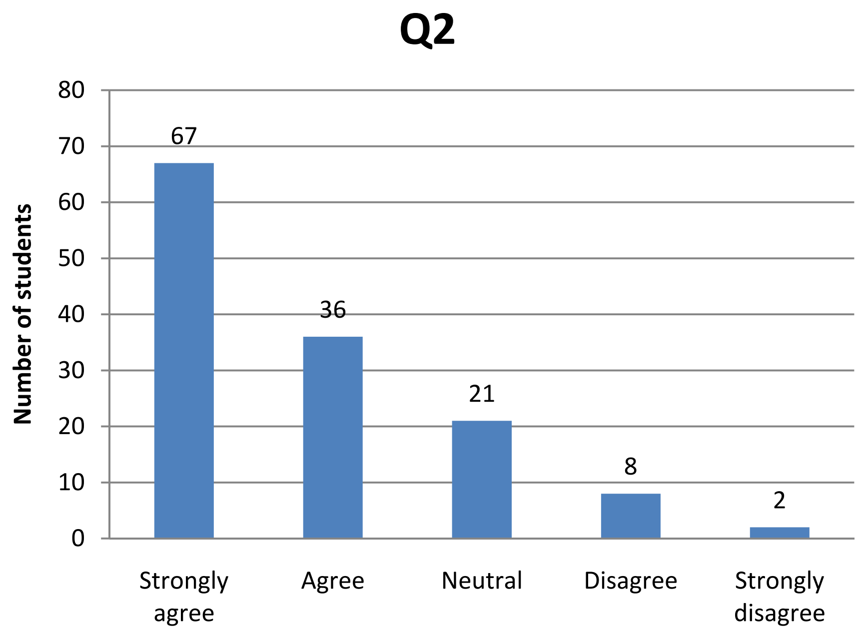
The engagement measurement is as follows. Sitting in pairs in the front and back sections of the lecture theatre, the trained observers would randomly select groups of 10-15 students that could be suitably observed. At five minute intervals, the observers would classify each student's behavior according to a list of engaged or disengaged behaviors (e.g. gesturing related to material, nodding in response to comment by instructor, text messaging, surfing web, reading unrelated book). If a student's behavior did not match one of the criteria, they were not counted, but this was a small fraction of the time. Measurements were not taken when students were voting on clicker questions because for some students this engagement could be too superficial to be meaningful as they were simply voting to get credit for responding to the question. Measurements were taken while students worked on the clicker questions when voting wasn't underway. This protocol has been shown by E. Lane and coworkers to have a high degree of inter-rater reliability after the brief training session of the observers.

2. Opinion survey and responses given in the experimental section

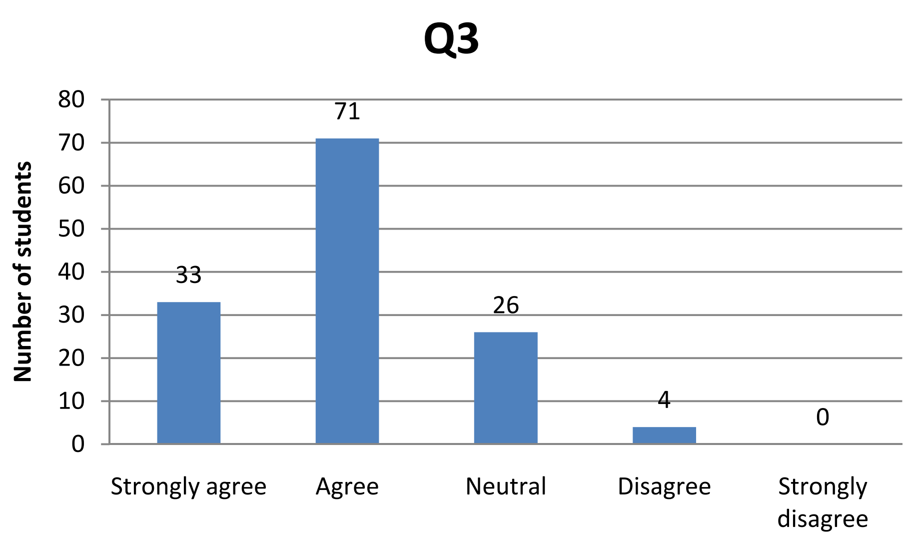
Q1 I really enjoyed the interactive teaching technique during the three lectures on E&M waves (Ch32):



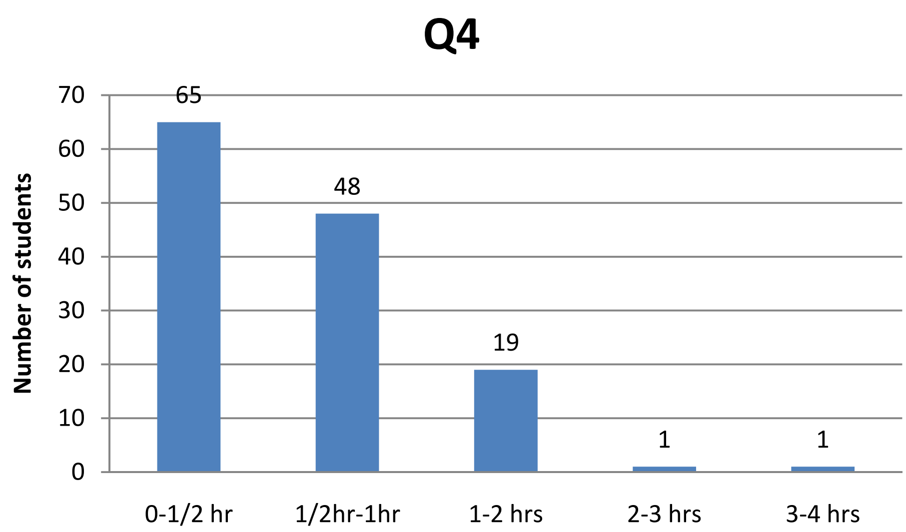
Q2 I feel I would have learned more if the whole course (Phys153) would have been taught in this highly interactive style:



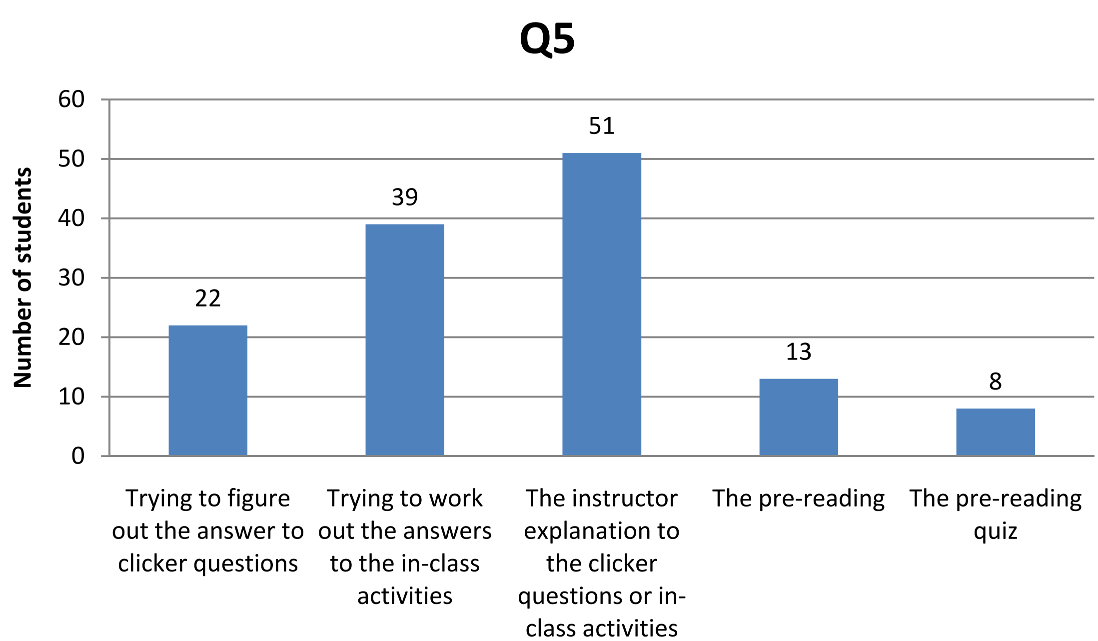
Q3 I thought the 30 min exam on E&M waves did a very good job at measuring how much I know about E&M waves and photons:



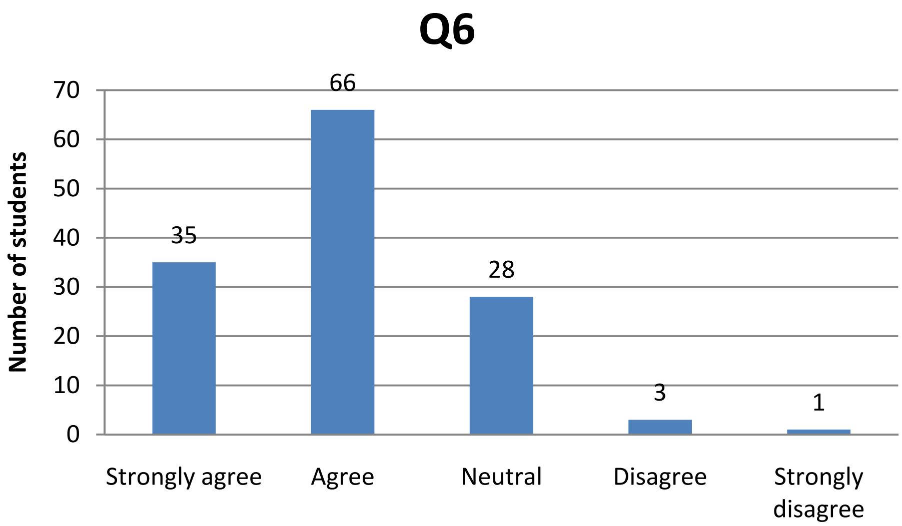
Q4 I studied for the E&M test/quiz for:



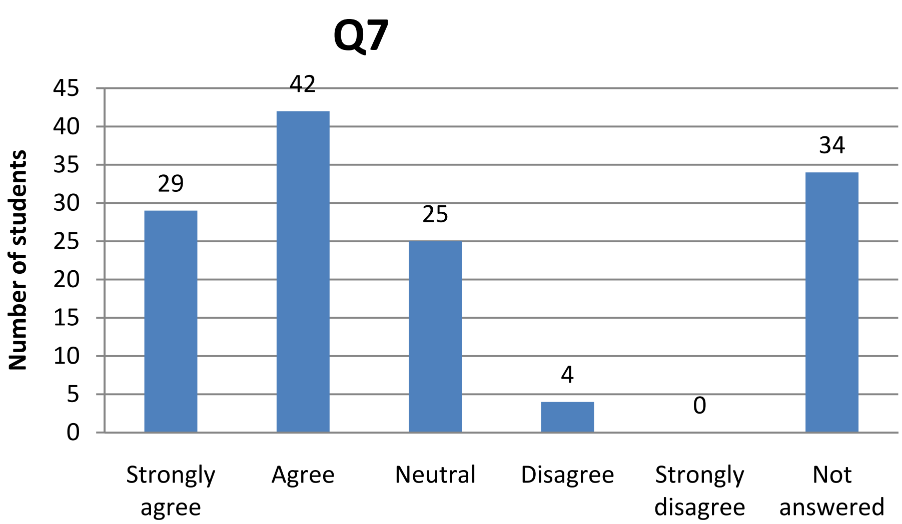
Q5 What contributed most to my learning during these three lecture on E&M waves:



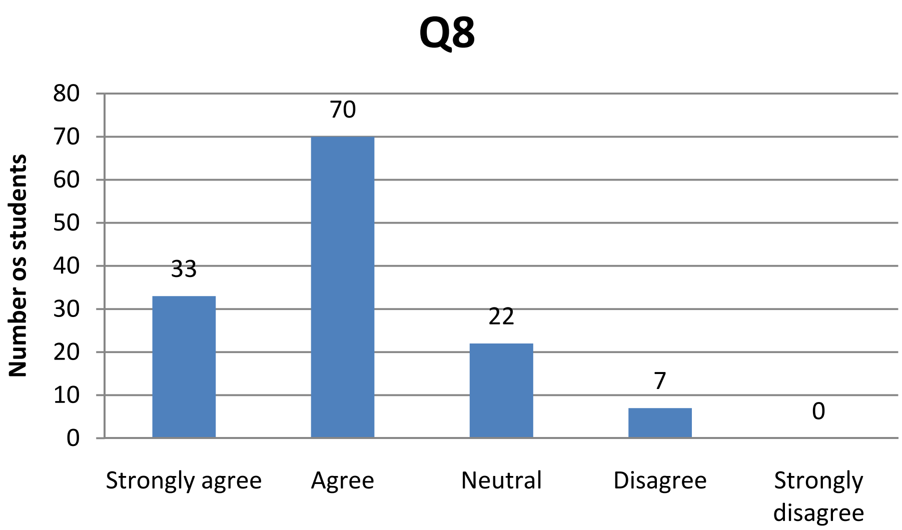
Q6 I found the pre-reading to be very helpful to my learning:



Q7 I found the pre-reading quiz to be very helpful to my learning:



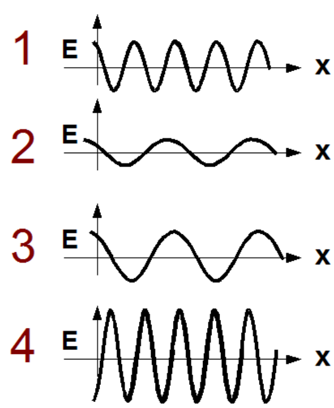
Q8 In class, the group discussions with my neighbors were very helpful to my learning:



3. <u>Test given to both sections on the material taught</u>

Question 1

The amplitude and frequency of 4 E&M waves are shown below. The waves are representative of one instant in time and are all travelling in vacuum. Which wave travels the fastest?



a) 1 c) 3
b) 2 d) 4 e) All 4 waves travel at same speed

Question 2

An Electromagnetic wave is traveling along the negative x-direction.
What is the direction of the Electric field vector E at a point where the Magnetic field vector B is in the positive y-direction?

(a) The E field points along the positive x-direction

(b) The E field points along the negative x-direction

(c) The E field points along the positive z-direction

(d) The E field points along the negative z-direction

(e) The E field points along the negative y-direction

Question 3

An electromagnetic wave is propagating along the positive x-direction with a magnetic field pointing along the z-direction:

$$B_z(x,t) = 3 \cdot 10^{-4} \sin(2\pi \cdot 10^6 x - \omega t)\ Tesla$$

What is the wavelength of this EM wave?
(*Note: 1 nanometer= 10^{-9} meter*)?

a) 10^4 nanometers

b) 10^3 nanometers

c) 100 nanometers

d) 10 nanometers

e) 1 nanometers

Question 4

An electromagnetic wave is propagating along the positive x-direction with a magnetic field pointing along the z-direction:

$$B_z(x,t) = 3 \cdot 10^{-4} \sin(2\pi \cdot 10^6 x - \omega t)\ Tesla$$

What is the strength of the Electric field ***E***?

a) 3×10^{-4} V/m
b) 9×10^{-8} V/m
c) 3 V/m
d) 9×10^{4} V/m
e) Not enough information is given

Question 5

An electromagnetic wave is propagating along the positive
x-direction with a magnetic field pointing along the z-direction:

$$B_z(x,t) = 3 \cdot 10^{-4} \sin(2\pi \cdot 10^{6} x - \omega t)\ Tesla$$

How will the intensity of the EM wave change if you increase the strength of the Magnetic field B_z by a factor of 4?

a) The intensity will increase by a factor of 16
b) The intensity will increase by a factor of 8
c) The intensity will increase by a factor of 4
d) The intensity will remain the same
e) Not enough information is given

Question 6

An electromagnetic wave is propagating along the positive x-direction with a magnetic field pointing along the z-direction:

$$B_z(x,t) = 3 \cdot 10^{-4} \sin(2\pi \cdot 10^{6} x - \omega t)\ Tesla$$

How will the intensity of the EM wave change if you decrease the *wavelength of the EM wave by a factor of 4*?

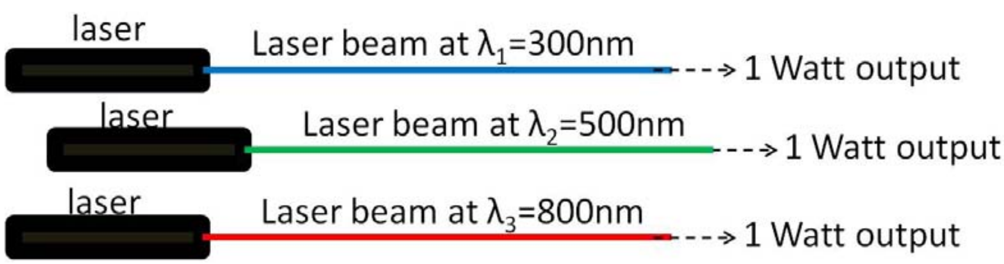
a) The intensity will decrease by a factor of 16
b) The intensity will increase by a factor of 16

c) The intensity will decrease by a factor of 4

d) The intensity will increase by a factor of 4

e) The intensity will remain the same

Question 7



Three laser beams have wavelengths λ_1=300nm, λ_2=500nm, and λ_3=800nm. The output power of all three lasers is **precisely 1Watt**.
Which laser emits the most energetic photons?

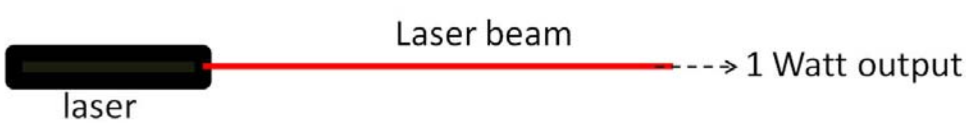
a) The Laser at λ_3=800nm

b) The Laser at λ_2=500nm

c) The Laser at λ_1=300nm

d) All three lasers emit photons with the same energy

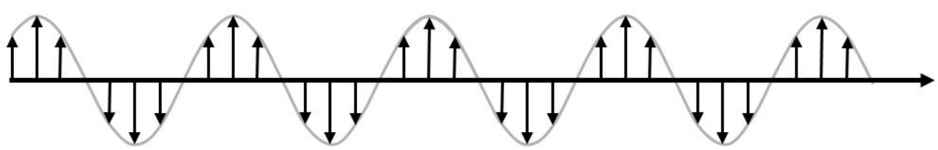
Question 8



The output wavelength of a laser is slowly changed from 450nm (Blue color) to 750nm (red color). While the wavelength is changed, the output power of the laser is kept precisely to 1Watt.
What can we say about the number of photons that are emitted by the laser every second?

a) Number of photons leaving the laser each second *decreases* as we increase λ
b) Number of photons leaving the laser each second *stays the same* as we increase λ
c) Number of photons leaving the laser each second *increases* as we increase λ
d) Not enough information is given

Question 9



$E(x,t) = E_{max}\sin(kx\text{-}wt)$ E_{max}=peak amplitude

What quantity best characterizes the energy/sec carried by the Electromagnetic wave?

a) frequency
b) wavelength (color)
c) E_{max}
d) $(E_{max})^2$
e) frequency^2

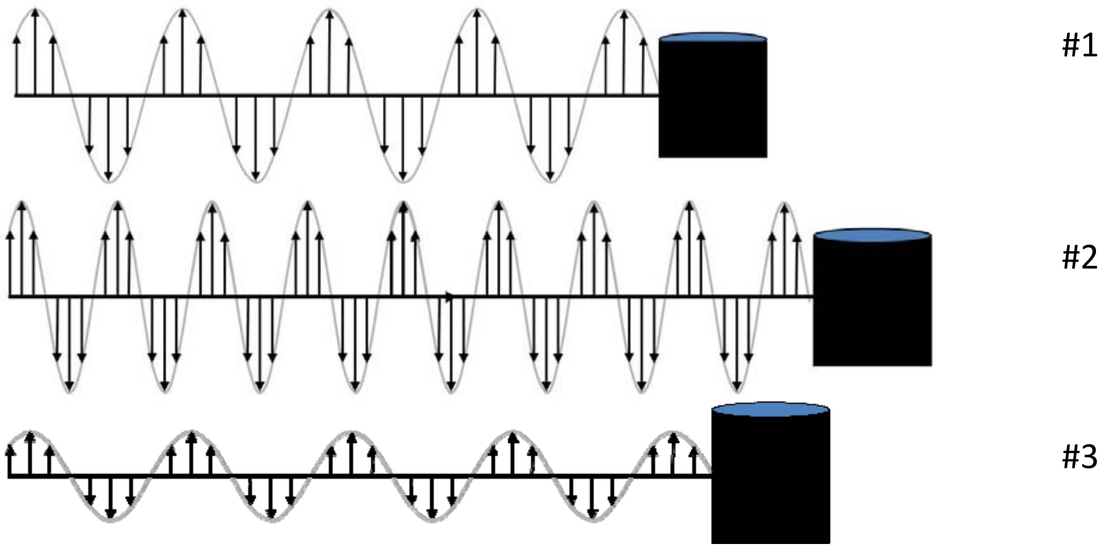
Question 10

True or False: In the absence of external forces, photons move along sinusoidal paths.

1) True
2) False

Question 11

3 Electromagnetic waves are absorbed by a dark object:



Which barrel will heat up the fastest?

a. 2>1>3 b. 1>2>3 c. 1=2>3
d. 1=3>2 e. 2>1=3

Where, E_{max}=peak

$\boxed{E_{1max}=E_{2max}>E_{3max}}$

Question 12

Light from the sun or from a light bulb appears to be constant (i.e. the rate at which the energy reaches your eyes doesn't appear to change in time). But we know that the strength of electromagnetic waves oscillates in time. So why do we see "steady" light? Pick the best answer.

a) The oscillations of the E and B fields cancel out so it looks like the rate of energy is constant
b) The oscillations in the rate of energy flow happen so quickly that we see an average energy which is steady
c) The maximum E and B fields are <u>constant</u>
d) You are looking over a large area so all the light combined will be constant

4. <u>In-class activities used in the experimental section for the three days</u>

The preparation of the in-class activities was based upon a "cognitive task analysis" of how physicists think about this material in terms of the mental models, multiple representations, related associations, and specific metacognitive processes they use with the different particular aspects of the material. The design of the activities also take into account known "naïve" student understandings or interpretations of particular aspects of this material that we were aware of from published literature or that LD and ES have observed in physics students. A full discussion of both these aspects is beyond the scope of this paper, but we have provided a brief annotation after each activity in italics to provide some guidance as to what expert-like thinking the activity is intended to stimulate the students to practice. This practice is primarily happening as the students formulate their answers and discuss the questions and answers with their fellow students and the instructor. As noted in the main text, the student questions and discussion often resulted in the coverage of material beyond what is shown in the activities presented here. There was also a few minute introduction to each class which is not reflected in the class notes shown here. We do not intend to imply that these activities are optimum. They were created by relatively inexperienced teachers as described in the main text, and with more experience with the course and the students these instructors could improve these activities.

The preparation of the experimental classes, which include class activities and reading quizzes, took roughly 20 person hours for the first class, dropping to 10 hours by the third class. Much of this preparation time was spent becoming familiar with the course material and, due to inexperience, designing activities for which there was not sufficient class time to utilize. The decrease in time required from the first to the third class is a reflection of increasing familiarity with the material and more experience with what these students could accomplish in a one hour class.

We estimate that under normal circumstances a moderately experienced instructor would require about 5hrs of preparation time per one hour class in this format. This includes: 3hrs to come up with clicker questions, activities, and reading quiz, 1hr of interview testing with one or two students, and 1hr to implement changes based on the student interview(s). Of course such material can be readily reused, in which case the preparation time would be far less.

Physics 153 Class Activities

CQ = Clicker Question
GT = Group Task

<u>Day 1</u>

CQ1

Which of the following is NOT one of Maxwell's Equations?

a) Gauss's Law for magnetism

b) $\frac{d^2E_y(x,t)}{dx^2} = \frac{1}{c^2}\frac{d^2E_y(x,t)}{dt^2}$

c) $\oint \vec{E} \cdot d\vec{l} = -\frac{d\Phi_B}{dt}$

d) Ampere's Law

Commentary: Largely factual review, but does practice expert distinction and relationship between Maxwell's equations and combination of Maxwell's equations that is the wave equation.

CQ2

Labelled 1-4 are Maxwell's equations in integral form. Labelled i-iv are the names of Maxwell's equations. Which of the following is the correct match?

1	$\oint \vec{E} \cdot d\vec{l} = -\frac{d\Phi_B}{dt}$	i) Ampere's Law
2	$\oiint \vec{E} \cdot d\vec{A} = \frac{Q_{enc}}{\epsilon_0}$	ii) Gauss's Law
3	$\oint \vec{B} \cdot d\vec{l} = \mu_0 i_{enc} - \frac{1}{c^2}\frac{d\Phi_E}{dt}$	iii) Gauss's Law for magnetism
4	$\oiint \vec{B} \cdot d\vec{A} = 0$	iv) Faraday's Law

a) 1i, 2ii, 3iv, 4iii
b) 1iv, 2ii, 3i, 4iii
c) 1ii, 2i, 3iii, 4iv
c) 1i, 2ii, 3iii, 4iv

Commentary: Factual memorization/review, not practicing expert thinking except small amount involved in translating between different mathematical representations.

CQ3

Which of the following best expresses what Gauss's Law describes?

a) The net electric flux through an enclosed surface is proportional to the net amount of charge inside the enclosed surface.
b) If you integrate over the electric field inside a box you get charge.
c) The net magnetic field along a closed path is proportional to the current flowing through the closed loop.
d) If you integrate the electric field over two parallel planar surfaces you get the charge enclosed between the two planar surfaces.

Commentary: Development of mental models of static electric and magnetic fields. Translation between representations, particularly between mathematical representation and physical models of electric and magnetic fields.

CQ4

Which of the following is true?

a) For EM waves to exist, they must propagate in a medium with atoms. With no atoms present, the field cannot have any effect on the system and therefore can't exist.
b) An EM wave can propagate through a vacuum.
c) An EM wave is like a wave travelling along a rope in that it needs atoms to move up and down.
d) An EM wave can *only* propagate in a vacuum since any medium would get in the way of its propagation.
e) More than one of the above is true.

Commentary: Develop and test mental model of EM wave. Practice metacognitive thinking in this context.

CQ5
Which of the following are forms of the wave equation for an EM wave propagating in vacuum along the x direction?

i) $$\frac{d^2 E_y(x,t)}{dx^2} = \epsilon_0 \mu_0 \frac{d^2 E_y(x,t)}{dt^2}$$

ii) $$\frac{dE_y(x,t)}{dx} = \epsilon_0 \mu_0 \frac{dE_y(x,t)}{dt}$$

iii) $$\frac{dB_z(x,t)}{dx} = \epsilon_0 \mu_0 \frac{dB_z(x,t)}{dt}$$

iv) $$\frac{d^2 B_z(x,t)}{dx^2} = \epsilon_0 \mu_0 \frac{d^2 B_z(x,t)}{dt^2}$$

a) i and iv
b) ii and iii
c) ii
d) i
e) None of the above

Commentary: Practicing translation between mathematical representations and physical phenomena.

GT
A friend of yours reminds you that en EM wave consists of both an E and B field.

She asks you if the following electric field
$E(x,t)=100x^2t$ Volts/m could be that of an EM wave.
Can you help? Be quantitative in your answer.

[*Hint*: Is there an equation that the electric field portion of an electromagnetic wave, E(x,t), must satisfy?]

Commentary: Recognize relationship between form of solution and its origin.

<u>Day 2</u>

CQ1

Which of the following are types of electromagnetic waves, just like the light coming from our sun?

a) FM radio (i.e. Signal picked up by your car)
b) Microwave (i.e. Popcorn)
c) Infrared (i.e. Night vision goggles)
d) X-rays (i.e. I just broke my leg)
e) all of the above

BONUS: Can you see with your eyes all EM radiation?

Commentary: Links to prior knowledge and building expert associations among previously encountered phenomena. Connect class material to real world phenomena.

CQ2

Could the following E wave function describe the electric field portion of a propagating EM wave?

$$E_y = E_{max} \cos{(kx)}$$

a) Yes
b) No
c) Not enough information to determine this

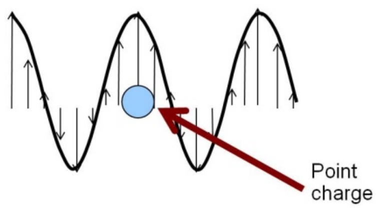
BONUS: What about cos(kxt)?
What about cos[k(x-vt)]?

Commentary: Translating between representations. Explicitly testing mathematical representations of physical phenomena.

GT

PhET Simulation: Radio Waves and Electromagnetic Fields
http://phet.colorado.edu/en/simulation/radio-waves
Observe the simulation of an EM wave being generated.

1. What do the arrows show?
2. A classmate tells you, "If I place a charge right there (see picture), the wave will pass over it and it won't affect it or apply a force on it". Do you agree with your classmate? Explain.



Commentary: Developing mental model, understand and apply expert representations and models to make predictions. Develop metacognitive capabilities.

CQ3
What is a source of EM waves?

a) A static charge distribution
b) A static current distribution
c) Charges moving at a constant speed
d) Accelerating charges
e) none of the above

Commentary: Developing and testing mental model, make explicit and provide feedback on known naïve interpretation.

CQ4
Someone has told you the maximum electric field strength and the electric field polarization of an electromagnetic wave. What do you know about the magnetic field?

i. Its maximum strength
ii. Its polarization
iii. Its propagation direction

a) i
b) i and ii
c) i and iii
d) ii and iii
e) all of the above
f) none of the above

Commentary: Sophisticated development and refinement of mental model, likely calling on multiple representations and self-checking in the process.

CQ5
Which of the following electromagnetic wave functions can describe a wave travelling in the negative y *direction*?

i)
$$\vec{E} = \hat{\imath}E_{max}\sin(ky + \omega t)$$
$$\vec{B} = \hat{k}B_{max}\sin(ky + \omega t)$$

iii)
$$\vec{E} = \hat{\jmath}E_{max}\sin(kx + \omega t)$$
$$\vec{B} = \hat{\jmath}B_{max}\sin(kx + \omega t)$$

ii)
$$\vec{E} = \hat{\imath}E_{max}\sin(ky + \omega t)$$
$$\vec{B} = -\hat{k}B_{max}\sin(ky + \omega t)$$

iv)
$$\vec{E} = \hat{\imath}E_{max}\sin(ky - \omega t)$$
$$\vec{B} = \hat{k}B_{max}\sin(ky - \omega t)$$

a) i b) iii c) iv d) i,ii and iv e) i and iv

Commentary: Translating between representations, relating mathematical representation to physical phenomena.

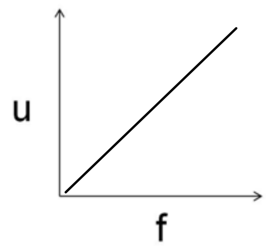
Day 3
CQ1
The frequency *f* of a laser pointer is increased but the light's *intensity* is unchanged. As a result, which of the following (perhaps more than one) are true? Explain.

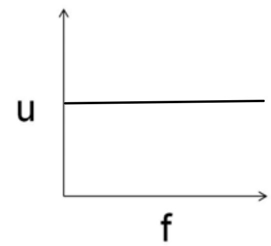
i) The output power is increased
ii) Each photon has more energy
iii) There are fewer photons per second
iv) There are more photons per second

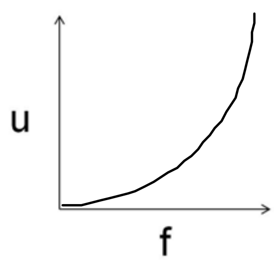
a) i
b) i and ii
c) ii and iii
d) ii and iv
e) iv

Commentary: Developing and testing mental models, building associations, confronting and providing targeted feedback on naïve understanding.

CQ2
Shown below are plots for the *energy density of an EM wave* vs. *frequency*. Think about how the energy density depends on the frequency of the wave. Which graph properly shows this relationship?

a) 

b) 

c) 

d) Not enough information to tell

e) None of the above

Commentary: Translating between representations, and in the process developing associations and refining mental model. Practicing metacognitive skill utilizing multiple representations.

CQ3
Many of you have learned in chemistry that photons are *quanta* of light. Which of the following best describes how photons and EM waves are related.

a) An EM wave is essentially made up of a single photon with frequency f; the size of which depends on the energy of the EM wave.
b) An EM wave is the sum of many photons that are all in phase.
c) An EM wave is composed of many photons where the strength of the wave depends on the energy of each photon and how many it is composed of.
d) The photons are what is moving up and down in an EM wave.
e) More than one statement is true

Commentary: Developing mental model by addressing prior knowledge and known naïve models.

GT

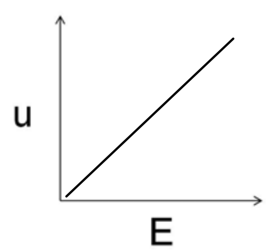
Three laser beams have wavelengths λ_1=400nm, λ_2=600nm and λ_3=800nm. The power (energy/sec) of each laser beam is the SAME at 1Watt. Rank in order, from largest to smallest:

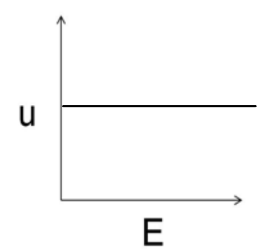
a) The photon energies $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$ in these three laser beams. Explain your answer.
b) The maximum strength of the E fields, E_{max1}, E_{max2}, E_{max3}, in these three laser beams. Explain your answer.
c) The number of photons per second N_1,N_2,N_3 delivered by the three laser beams.

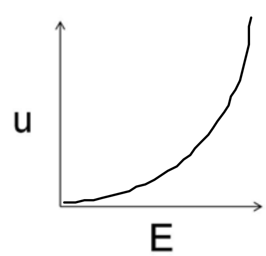
Commentary: A transfer task requiring recognition of relevant variables and use of mental model.

CQ4

Shown below are plots of energy density vs. electric field strength for an EM wave. Think about how the energy density depends on the electric field strength. Which graph properly shows this relationship?

a)


b)


c)


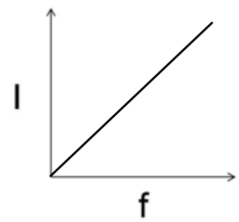
d) Not enough information to tell

e) None of the above

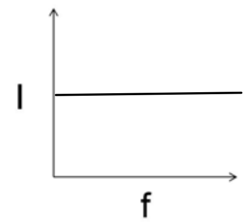
Commentary: Similar to CQ2

CQ5
Shown below are plots of intensity vs. frequency for a classical EM wave. Think about how the intensity depends on the frequency of the wave. Which graph properly shows this relationship?

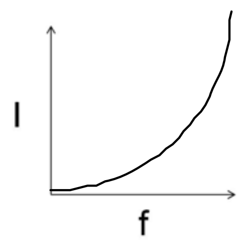
a)



b)



c)



d) Not enough information to tell

e) None of the above

Commentary: Similar to CQ2, and addressing and providing feedback to correct known naïve thinking.

5. <u>Learning objectives agreed upon by the two instructors</u>

The learning objectives were categorized into levels of importance with A being the most important to C being less important. The test primarily covered the category A objectives. Although we believe it would be educationally beneficial to provide the students with such objectives in class before the unit, in deference to the wishes of the instructor of the control section, the students were not given the learning objectives.

After completing this module on EM waves the students should:

A

1) Be able to write down the wave equation for electric and magnetic fields.
2) Be able to describe the characteristics of a plane wave.
    a) Direction of propagation
    b) Polarization
    c) Planes of constant phase (C)
3) Be able to write the relationships between wavespeed, wavelength, frequency,

angular frequency and wave vector.

4)
   a) Given an analytical expression for an EM wavefunction (E or B), be able to represent it graphically
   b) Be able to correctly interpret all of the features of the representation when plotted as a function of time or space, *i.e*. Amplitude corresponds to field strength, being able to identify wavelength, frequency etc (see 3)
5) Be able to write down the relationship between polarizations of the E and B fields of an EM wave and its direction of propagation.
6) Be able to identify the equation of energy density of an EM waves in terms of E and B and in terms of just E, *i.e*. know that it goes at E^2 and doesn't depend on frequency
7) Be able to contrast EM waves with mechanical waves
   a) Compare how energy depends on critical parameters such as amplitude and frequency
   b) Compare physical interpretation of their oscillating amplitude
   c) Appreciate the fact that EM waves propagate in a vacuum.
8) Be able to give a basic description of how EM waves are related to photons.
   a) Be able to contrast the energy dependence on critical parameters for EM waves and photons.
   b) For an EM wave with a given intensity, be able to identify how many photons of a given frequency it is composed of.
9)
   a) Be able to write the Poynting vector in terms of E and B.
   b) Be able to describe how the intensity is related to the Poynting vector
   c) Be able to give a basic description of what the Poynting vector represents.

B

1) Be able to identify Maxwell's Equations by name.
2) Be able to test whether **scalar** E and B wavefunctions for an electromagnetic wave satisfy the wave equation.
   a) Given an E and B equation plug it into the wave equation and check that the sides of the equation equate.
3) Be able to identify a set of **vector** E and B wave functions that properly describe an EM wave propagating in a given direction.
   a) Use right hand rule

C

1) Qualitatively be able to explain the meaning of Gauss' Laws, Faradays Law and Ampere's Law
2) Be able to identify the terms in Maxwell's Equations that lead to the wave equation (*i.e*. Plane wave light propagation)

3) Be able to give examples of transverse waves. Contrast transverse waves with longitudinal waves.
4) Give examples of how we experience the energy of EM waves in everyday life.
   a) Ex. From the sun get: heat, can power solar cells etc
   b) Ex. Need batteries to power flashlight
   c) Etc.
5) Be able to identify points of equal phase along on a wave.

6. Hawthorne effect discussion

It is not the intention of this paper to review the Hawthorne effect and its history, but we comment on it only because this is such a frequent question raised about this work. It is not plausible that it resulted in a significant impact on the results reported here. As discussed extensively in (*S1-S3*), analyses of the methodology and data used in the original Hawthorne plant studies reveal both serious flaws in the methodology, and an absence of statistically significant data supporting the existence of the claimed effect. Thus, the failure to replicate such an effect in an educational setting, as reported in (*S4*), is not surprising.

Even if the Hawthorne effect were true, namely that people engaged in routine tasks will improve performance when conditions are changed in any manner, it would not be very relevant to this experiment. If one examines the typical daily activities of these students, the differences introduced by this experiment are not a significant increase in the variety of their educational experiences. These students are going to a variety of classes every day. These classes incorporate both a wide variety of subjects and instructional styles. They have large and small lecture courses, seminar courses, instructional labs, recitation sections, and project lab courses, all with various types of individual and group assignments. So while this experiment is introducing change in the student experience in one particular course (3 total hours per week) it provides little incremental novelty to their overall daily educational experience.

Finally, there have been several other full length physics courses at UBC transformed following the same design as discussed here. Those courses had much higher attendance and engagement for the entire term than is typical for other UBC physics courses including previous offerings of those courses. The attendance was similar or higher than what was observed in the experimental section in this work, and the engagement appeared to be similar. There were no control groups for those courses that can be used for learning comparisons however. This indicates that the level of attendance and engagement reported here were due to the instructional design and not merely due to the one week novelty.

7. List of proven teaching practices used

The instruction in this experiment incorporates variants on many established active learning instructional techniques. These include Just In Time Teaching (*S5*), Peer Instruction (*S6)*, some elements of Scale Up (*S7*), use of clicker question practices to facilitate student thinking and effective feedback as discussed in (*S8*) and (*S9*), some elements of Interactive Lecture Demonstrations (*S10*), group work (*S11*) and numerous other references), and the use of interactive simulations (*S12*). See also (*S13*) for a more extensive set of references on these teaching practices.

Supplemental references

S1. B. Rice, The Hawthorne defect: Persistence of a flawed theory. *Psychology Today* **16**, 70-74 (1982).

S2. S. R. G. Jones, Was there a Hawthorne effect? *Am. J. Sociology* **98**, 451-468 (1992).

S3. H. M. Parsons, What happened at Hawthorne? *Science* **183**, 922-932 (1974).

S4. R. H. Bauernfeind, C. J. Olson, Is the Hawthorne effect in Educational Experiments a Chimera? *Phi Delta Kappan* **55**, 271-273 (1973).

S5. G. M. Novak, E. T. Patterson, A. D. Gavrin, W. Christian, *Just-in-Time Teaching: Blending Active Learning with Web Technology* (Prentice Hall, New Jersey, 1999).

S6. C. H. Crouch, E. Mazur, Peer instruction: Ten years of experience and results. *Am. J. Phys*. **69**, 970-977 (2001).

S7. R. J. Beichner et al., "The Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) Project," in *Research-Based Reform of University Physics*, E. F. Redish, P. J. Cooney, Eds. (American Association of Physics Teachers, College Park, MD, 2007).

S8. I. D. Beatty, W. J. Gerace, W. J. Leonard, R. J. Dufresne, Designing effective questions for classroom response system teaching. *Am. J. Phys*. **74**, 31-39 (2006).

S9. University of Colorado Science Education Initiative & University of British Columbia Carl Wieman Science Education Initiative, *Clicker resource guide: An instructor's guide to the effective use of personal response systems (clickers) in teaching* (2008). Available at http://www.cwsei.ubc.ca/resources/files/Clicker_guide_CWSEI_CU-SEI.pdf.

S10. D. R. Sokoloff, R. K. Thornton, *Interactive Lecture Demonstrations, Active Learning in Introductory Physics* (John Wiley & Sons, Inc, New Jersey, 1994).

S11. P. Heller, R. Keith, S. Anderson, Teaching Problem Solving Through Cooperative Grouping. Part 1: Group versus Individual Problem Solving. *Am. J. Phys.* **60**, 628-636 (1992).

S12. C. E. Wieman, W.K. Adams, K.K. Perkins, PhET: Simulations that Enhance

Learning. *Science* **322**, 682-683 (2008) and http://phet.colorado.edu.

S13. J. E. Froyd, "Evidence for the Efficacy of Student-active Learning Pedagogies" (Project Kaleidoscope, 2007). Available at http://www.pkal.org/documents/BibliographyofSALPedagogies.cfm.

Compiled by Clarissa Dirks, 2011

Evidence for the Efficacy of Active Learning in the Life Sciences

References

- Anderson, W.L., S.M. Mitchell, and M.P. Osgood, Comparison of Student Performance in Cooperative Learning and Traditional Lecture-Based Biochemistry Classes. Biochemistry and Molecular Biology Education., 2005.
- Armbruster, P., et al., Active Learning and Student-Centered Pedagogy Improve Student Attitudes and Performance in Introductory Biology. CBE - Life Sciences Education., 2009.
- Armstrong, N., S.-M. Chang, and M. Brickman, Cooperative Learning in Industrial-Sized Biology Classes. CBE - Life Sciences Education., 2007.
- Born, W.K., W. Revelle, and L.H. Pinto, Improving Biology Performance with Workshop Groups. Journal of Science Education and Technology., 2002.
- Brauner, A., et al., Open-Ended Assignments and Student Responsibility. Biochemistry and Molecular Biology Education., 2007.
- Brickman, P., Gormally, C., Armstrong, N. and Hallar, B., Effects of Inquiry-based Learning on Students’ Science Literacy Skills and Confidence. International Journal for the Scholarship of Teaching and Learning, 2009. 3(2).
- Burrowes, P. and G. Nazario, Promoting Student Learning through the Integration of Lab and Lecture: The Seamless Biology Curriculum. Journal of College Science Teaching., 2008.
- Burrowes, P.A., A Student-Centered Approach to Teaching General Biology That Really Works: Lord's Constructivist Model Put to a Test. American Biology Teacher., 2003.
- Butler, J.A., Use of Teaching Methods within the Lecture Format. Medical Teacher., 1992.
- Buxeda, R.J. and D.A. Moore, Transforming a sequence of microbiology courses using student profile data. Microbiology Education, 2000. 1(1): p. 1-6.
- Caldwell, J.E., Clickers in the Large Classroom: Current Research and Best-Practice Tips. CBE - Life Sciences Education., 2007.
- Casem, M.L., Active Learning Is Not Enough. Journal of College Science Teaching., 2006.
- Casem, M.L., Student Perspectives on Curricular Change: Lessons from an Undergraduate Lower-Division Biology Core. CBE - Life Sciences Education., 2006.
- Christianson, R.G. and K.M. Fisher, Comparison of Student Learning about Diffusion and Osmosis in Constructivist and Traditional Classrooms. International Journal of Science Education., 1999.
- Cliff, W.H. and A.W. Wright, Directed case study method for teaching anatomy and physiology. Advances in Physiology Education, 1996. 15(1): p. 19-28.
- Cortright, R.N., et al., Student retention of course content is improved by collaborative-group testing. Advances in Physiology Education, 2003. 27: p. 102-108.
- Cortright, R.N., H.L. Collins, and S.E. DiCarlo, Peer Instruction Enhanced Meaningful Learning: Ability to Solve Novel Problems. Advances in Physiology Education., 2005.
- Daempfle, P.A., Instructional Approaches for the Improvement of Reasoning in Introductory College Biology Courses: A Review of the Research. 2002.
- D'Avanzo, C., Biology Concept Inventories: Overview, Status, and Next Steps. BioScience., 2008. 39(11): p. 952-978.
- Dirks, C. and M. Cunningham, Enhancing Diversity in Science: Is Teaching Science Process Skills the Answer? CBE - Life Sciences Education., 2006. 5(3): p. 218-226.
- Ebert-May, D., C. Brewer, and S. Allred, Innovation in Large Lectures--Teaching for Active Learning. BioScience., 1997.

- Fonseca, A., C.I. Extremina, and A. Freitas da Fonseca. Concept Mapping: A Strategy for Meaningful Learning in Medical Microbiology. in First International Conference on Concept Mapping. 2004. Pamplona Spain.
- Francom, G., et al., Biology 100: A Task-Centered, Peer-Interactive Redesign. TechTrends: Linking Research and Practice to Improve Learning., 2009.
- Freeman, S., et al., Prescribed Active Learning Increases Performance in Introductory Biology. CBE - Life Sciences Education., 2007.
- Froyd, J., Evidence for the Efficacy of Student-Active Learning Pedagogies. 2008, Texas A&M University: College Station. p. 1-20.
- Gehring, K.M. and D.A. Eastman, Information Fluency for Undergraduate Biology Majors: Applications of Inquiry-based Learning in a Developmental Biology Course. CBE - Life Sciences Education., 2008.
- Goodwin, L., J.E. Miller, and R.D. Cheetham, Teaching Freshman to Think: Does Active Learning Work? BioScience., 1991. 41(10): p. 719-722.
- Halme, D.G., et al., A Small-Scale Concept-Based Laboratory Component: The Best of Both Worlds. CBE - Life Sciences Education., 2006.
- Harris, M.A., et al., A Combination of Hand-Held Models and Computer Imaging Programs Helps Students Answer Oral Questions about Molecular Structure and Function: A Controlled Investigation of Student Learning. CBE - Life Sciences Education., 2009.
- Hoffman, E.A., Successful Application of Active Learning Techniques to Introductory Microbiology. Microbiology Education, 2001. 2(1): p. 5-11.
- Hoffman, K., et al., Problem-Based Learning Outcomes: Ten Years of Experience at the University of Missouri-Columbia School of Medicine. Academic Medicine, 2006. 81(7): p. 617-625.
- Hufford, T.L., Increasing Academic Performance in an Introductory Biology Course. BioScience., 1991.
- Huxham, M., Learning in Lectures: Do 'Interactive Windows' Help? Active Learning in Higher Education., 2005.
- Johnson, M.A. and A.E. Lawson, What Are the Relative Effects of Reasoning Ability and Prior Knowledge on Biology Achievement in Expository and Inquiry Classes? Journal of Research in Science Teaching., 1998.
- Kitchen, E., et al., Teaching Cell Biology in the Large-Enrollment Classroom: Methods to Promote Analytical Thinking and Assessment of Their Effectiveness. Cell Biology Education., 2003.
- Klappa, P., Promoting Active Learning through "Pub Quizzes"--A Case Study at the University of Kent. Bioscience Education., 2009.
- Klionsky, D.J., Constructing Knowledge in the Lecture Hall. Journal of College Science Teaching., 2002.
- Knight, J.K. and W.B. Wood, Teaching More by Lecturing Less. Cell Biology Education., 2005.
- Lawson, A., et al., Evaluating College Science and Mathematics Instruction: A Reform Effort That Improves Teaching Skills. Journal of College Science Teaching., 2002.
- Libarkin, J., Concept Inventories in Higher Education Science, in A manuscript prepared for the National Research Council Promising Practices in Undergraduate STEM Education Workshop 2. 2008: Washington, D.C.
- Lord, T.R. Is the Final Grade in College Biology a True Measure of Student Knowledge? 1988.
- Marbach-Ad, G. and P.G. Sokolove, Can Undergraduate Biology Students Learn To Ask Higher Level Questions? Journal of Research in Science Teaching., 2000.

- Marbach-Ad, G. and P.G. Sokolove, The Use of E-Mail and In-Class Writing To Facilitate Student-Instructor Interaction in Large-Enrollment Traditional and Active Learning Classes. Journal of Science Education and Technology., 2002.
- Marrs, K.A. and G.W. Chism, III, Just-in-Time Teaching for Food Science: Creating an Active Learner Classroom. Journal of Food Science Education., 2005.
- McCrindle, A.R. and C.A. Christensen, The Impact of Learning Journals on Metacognitive and Cognitive Processes and Learning Performance. Learning and Instruction., 1995.
- McDaniel, C.N., et al., Increased Learning Observed in Redesigned Introductory Biology Course that Employed Web-Enhanced, Interactive Pedagogy. CBE - Life Sciences Education., 2007.
- McInerney, M.J. and L.D. Fink, Team-Based learning enhances long-term retention and critical thinking in an undergraduate microbial physiology course Microbiology Education, 2003. 4: p. 3-12.
- Michael, J., Where's the evidence that active learning works? Advances in Physiology Education, 2006. 30: p. 159-167.
- Montelone, B.A., D.A. Rintoul, and L.G. Williams, Assessment of the Effectiveness of the Studio Format in Introductory Undergraduate Biology. CBE - Life Sciences Education., 2008.
- Nehm, R. and L. Reilly, Biology majors' knowledge and misconceptions of natural selection. BioScience., 2007. 57(3): p. 263-272.
- Nelson, J., et al., Cloning the Professor, an Alternative to Ineffective Teaching in a Large Course. CBE - Life Sciences Education., 2009.
- Orsmond, P., S. Merry, and A. Callaghan, Implementation of a Formative Assessment Model Incorporating Peer and Self-Assessment. Innovations in Education and Teaching International., 2004.
- Orsmond, P., S. Merry, and K. Reiling, A Study in Self-Assessment: Tutor and Students' Perceptions of Performance Criteria. Assessment and Evaluation in Higher Education., 1997.
- Orsmond, P., S. Merry, and K. Reiling, The Use of Student Derived Marking Criteria in Peer and Self-Assessment. Assessment and Evaluation in Higher Education., 2000.
- Osborne, J., Arguing to learn in science: the role of collaborative, critical discourse. Science, 2010. 328(5977): p. 463-6.
- Paschal, C.B., Formative assessment in physiology teaching using a wire- less classroom communication system. Advances in Physiology Education, 2002. 26: p. 299-308.
- Pelaez, N., Problem-based writing with peer review improves academic performance in physiology. Advances in Physiology Education, 2002. 26: p. 174-184.
- Perez, K.E., et al., Does Displaying the Class Results Affect Student Discussion during Peer Instruction? CBE - Life Sciences Education., 2010.
- Peters, A., Teaching biochemistry at a minority-serving institution: An evaluation of the role of collaborative learning as a tool for science mastery. Journal of Chemical Education, 2005. 82(4): p. 571-574.
- Pisani, A.M. Involvement through Cooperative Learning: An Attempt To Increase Persistence in the Biological Sciences. ASHE Annual Meeting Paper. 1994.
- Preszler, R.W., Cooperative Concept Mapping: Improving Performance in Undergraduate Biology. Journal of College Science Teaching, 2004. 33(6): p. 30-35.
- Preszler, R.W., et al., Assessment of the Effects of Student Response Systems on Student Learning and Attitudes over a Broad Range of Biology Courses. CBE - Life Sciences Education., 2007. 6(1): p. 29-41.

- Preszler, R.W., Replacing Lecture with Peer-Led Workshops Improves Student Learning. CBE - Life Sciences Education., 2009.
- Preszler, R.W., Student and teacher learning in a supplemental learning biology course. Bioscene: Journal of College Biology Teaching., 2006. 32(2): p. 21-25.
- Rao, S.P., H.L. Collins, and S.E. DiCarlo, Collaborative testing enhances student learning. Advances in Physiology Education, 2002. 26: p. 37-41.
- Roy, H., Studio vs. Interactive Lecture Demonstration--Effects on Student Learning. Bioscene: Journal of College Biology Teaching., 2003. 29(1): p. 3-6.
- Rybarczyk, B.J., et al., A Case-based approach increases student learning outcomes and comprehension of cellular respiration concepts. Biochemistry and Molecular Biology Education., 2007. 35(3): p. 181-186.
- Schroeder, C.M., et al., A Meta-Analysis of National Research: Effects of Teaching Strategies on Student Achievement in Science in the United States. Journal of Research in Science Teaching., 2007.
- Smith, M.K., et al., Why peer discussion improves student performance on in-class concept questions. Science, 2009. 323(5910): p. 122-4.
- Steele, J., Effect of essay-style lectures quizzes on student performance on anatomy and physiology exams. Bioscene: Journal of College Biology Teaching., 2003. 29(4): p. 15-20.
- Tessier, J., Small-group peer teaching in an introductory biology classroom. Journal of College Science Teaching, 2007. 36(4): p. 64-69.
- Tessier, J., Using peer teaching to promote learning in biology. Journal of College Science Teaching, 2004. 33(6): p. 16-19.
- Traver, H.A., et al., Student Reactions and Learning: Evaluation of a Biochemistry Course That Uses Web Technology and Student Collaboration. Biochemistry and Molecular Biology Education., 2001.
- Trempey, J.E., M.M. Skinner, and W.A. Siebold, Learning microbiology through cooperation: Designing cooperative learning activities that promote interdependence, interaction, and accountability. Microbiology Education, 2002. 3(1): p. 26-36.
- Udovic, D., et al., Workshop Biology: Demonstrating the Effectiveness of Active Learning in an Introductory Biology Course. Bioscience., 2002.
- Walker, J.D., et al., A Delicate Balance: Integrating Active Learning into a Large Lecture Course. CBE - Life Sciences Education., 2008.
- Wilke, R.R., The effect of active learning on student characteristics in a human physiology course for nonmajors. Advances in Physiology Education, 2003. 27: p. 207-223.
- Windschitl, M. Student Epistemological Beliefs and Conceptual Change Activities: How Do Pair Members Affect Each Other? 1997.
- Wood, W.B., Innovations in teaching undergraduate biology and why we need them. Annual Review of Cell and Developmental Biology, 2009. 25: p. 93-112.
- Wright, R., Learning cell biology as a team: A project-based approach to upper-division cell biology. Cell Biology Education, 2002. 1: p. 145-153.
- Yarden, H., G. Marbach-Ad, and J.M. Gershoni, Using the Concept Map Technique in Teaching Introductory Cell Biology to College Freshmen. Bioscene: Journal of College Biology Teaching., 2004.

Evidence for the Efficacy of Student-active Learning Pedagogies

Jeffrey E. Froyd
Research Professor
Director of Faculty and Organizational Development
Texas A&M University
froyd@tamu.edu
Version: 5 August 2008

Summary

Although many resources have been published on improvements in student retention and/or learning as a result of using what can be referred to as student-active pedagogies, the resources are published in a variety of journals or on various websites. As a result, it may be difficult for an individual to locate and assemble these resources to support an argument in favor of using these alternative pedagogies. Over a period of eight years, including my time as the Project Director for the Foundation Coalition, one of the Engineering Education Coalitions supported by NSF, I have tried to assemble many of these resources in one place for easy reference.

Cooperative and Small-group Learning Pedagogies

Springer, L., Stanne, M. E., and Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of Educational Research, 69*(1), 21–51.

Abstract: Recent calls for instructional innovation in undergraduate science, mathematics, engineering, and technology (SMET) courses and programs highlight the need for a solid foundation of educational research at the undergraduate level on which to base policy and practice. We report herein the results of a meta-analysis that integrates research on undergraduate SMET education since 1980. The meta-analysis demonstrates that various forms of small-group learning are effective in promoting greater academic achievement, more favorable attitudes toward learning, and increased persistence through SMET courses and programs. The magnitude of the effects reported in this study exceeds most findings in comparable reviews of research on educational innovation and supports more widespread implementation of small-group learning in undergraduate SMET.

Quote: The 0.51 effect of small-group learning on achievement reported in this study would move a student from the 50th percentile to the 70th on a standardized (norm-referenced) test. Similarly, a 0.46 effect size on students' persistence is enough to reduce attrition from SMET courses by 22%.

Hake, R. R. (1998). Interactive-Engagement vs. Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses. *American Journal of Physics, 66*(1), 64–74

Abstract: A survey of pre/post test data using the Halloun-Hestenes Mechanics Diagnostic test or more recent Force Concept Inventory is reported for 62 introductory physics courses enrolling a total number of students N = 6542. A consistent analysis over diverse student populations in high schools, colleges, and universities is obtained if a rough measure of the average effectiveness of a course in promoting conceptual understanding is taken to be the average normalized gain <g>. The latter is defined as the ratio of the actual average gain (%<post> – %<pre>) to the maximum possible average gain (100–%<pre>). Fourteen "traditional" (T) courses (N = 2084) which made little or no use of interactive-engagement (IE) methods achieved an average gain <g>$_{\text{T-ave}}$ = 0.23 ± 0.04 (std dev). In sharp contrast, forty-eight courses (N = 4458) which made substantial use of IE methods achieved an average gain <g>$_{\text{IE-ave}}$ = 0.48 ± 0.14 (std dev), almost two standard deviations of <g>$_{\text{IE-ave}}$ above that of the traditional courses. Results for 30 (N = 3259) of the above 62 courses on the problem-solving Mechanics Baseline test of Hestenes-Wells imply that IE strategies enhance problem-solving ability. The conceptual and problem-solving test results strongly suggest that the classroom use of IE methods can increase mechanics-course effectiveness well beyond that obtained in traditional practice

Wage, K. E., Buck, J. R., Wright, C. H. G., and Welch, T. B. (2005). The Signals and Systems Concept Inventory. *IEEE Transactions on Education, 48*(3), 448–461

Abstract: The signal processing community needs quantitative standardized tools to assess student learning in order to improve teaching methods and satisfy accreditation requirements. The Signals and Systems Concept Inventory (SSCI) is a 25-question multiple-choice exam designed to measure students' understanding of fundamental concepts taught in standard signals and systems curricula. When administered as a pre- and postcourse assessment, the SSCI measures the gain in conceptual understanding as a result of instruction. This paper summarizes the three-year development of this new assessment instrument and presents results obtained from testing with a pool of over 900 students from seven schools. Initial findings from the SSCI study show that students in traditional lecture courses master approximately 20% of the concepts they do not know prior to the start of the course. Other results highlight the most common student misconceptions and quantify the correlation between signals and systems and prerequisite courses.

Buck, J. R., and Wage, K. E. (2005). Active and Cooperative Learning in Signal Processing Courses. *IEEE Signal Processing Magazine, 22*(2), 76–81

Abstract: None

Quote: Pedagogical research in physics has found that (g) is robust to variations in instructor experience, student background, class size, and university ranking [2], [3]. Hake's major conclusion was that 14 traditional lecture format classes achieved normalized gain (g) = 0.23 ± 0.04, while 48 IE (or ACL) courses achieved (g) = 0.48 ± 0.14, nearly two standard deviations better than lecture courses. Subsequent papers have reported similar performance for IE methods in physics courses [4]. In our study using the SSCI, we found results strikingly similar to those reported by Hake. We computed (g) for 20 signals and systems courses. The 15 lecture format courses had normalized gain (g) = 0.20 ± 0.07, while the five ACL courses for which we have data achieved (g) = 0.37 ± 0.06. The gain for these ACL courses is more than two standard deviations above the lecture courses.

Crouch, C.H., and Mazur, E. (2001). *Peer Instruction: Ten years of experience and results. American Journal of Physics, 69*(9), 970–977

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

Wright, J.C., Millar, S.B., Kosciuk, S.A., Penberthy, D. L., Williams, P.H., and Wampold, B.E. (1998). A Novel Strategy for Assessing the Effects of Curriculum Reform on Student Competence. *Journal of Chemical Education, 85*(8), 986–992

Abstract: A new strategy has been developed to credibly assess the effects of curriculum reform on student competence. In order to implement the strategy, a comparative assessment was performed between the students in a section of a course with active learning and those in a reference section. The comparison used 25 faculty to conduct oral interviews that assessed student competence using each faculty member's definition of competence. Qualitative research methods were also employed to identify the reasons for any differences. The results show substantial differences in the students' reasoning and self expression skills that we believe are directly attributable to their structured active learning experiences.

Quote: Although SAL [structured active learning] students outperformed RL [responsive lecturing] students in all subcategories, the assessors in the meta-awareness subgroup found the largest differences between sections—almost ½ the maximum possible differences. [This difference is statistically significant.] This finding indicates that the major reason for the large difference in student competence was the thinking process that students displayed during the oral examination. In the analysis and agility subgroups, the differences became smaller but were still 25% of the maximum possible difference. The differences are not significant, however, because of the small number of assessors in each category. In the analogy subgroup, the differences are still smaller and are not significant

Prince, M. (2004). Does Active Learning Work? A Review of the Research. *Journal of Engineering Education, 93*(3), 223–231

Abstract: This study examines the evidence for the effectiveness of active learning. It defines the common forms of active learning most relevant for engineering faculty and critically examines the core element of each method. It is found that there is broad but uneven support for the core elements of active, collaborative, cooperative and problem-based learning.

Quote: The reported results are consistently positive. Indeed, looking at high quality studies with good internal validity, the already large effect size of 0.67 shown in Table 2 for academic achievement increases to 0.88. In real terms, this would increase a student's exam score from 75 to 85 in the "classic" example cited previously, though of course this specific result is dependent on the assumed grade distribution. As seen in Table 2, cooperation also promotes interpersonal relationships, improves social support and fosters self-esteem.

Quote: In summary, there is broad empirical support for the central premise of cooperative learning, that cooperation is more effective than competition for promoting a range of positive learning outcomes. These results include enhanced academic achievement and a number of attitudinal outcomes. In addition, cooperative learning provides a natural environment in which to enhance interpersonal skills and there are rational arguments and evidence to show the effectiveness of cooperation in this regard.

Johnson, D. W., Johnson, R. T., and Smith, K. A. (1998). Cooperative Learning Returns to College: What Evidence Is There That It Works? *Change*, 30(4), 26–35

Abstract: None

Quote: Between 1924 and 1997, over 168 studies were conducted comparing the relative efficacy of cooperative, competitive, and individualistic learning on the achievement of individuals 18 years or older. These studies indicate that cooperative learning promotes higher individual achievement than do competitive approaches (effect size = 0.49) or individualistic ones (effect size = 0.53). Effect sizes of this order describe significant, substantial increases in achievement. They mean, for example, that college students who would score at the 50th percentile level when learning competitively will score in the 69th percentile when learning cooperatively; students who would score at the 53rd percentile level when learning individualistic ally will score at the 70th percentile when learning cooperatively.

Bowen, C. W. (2000). A Quantitative Literature Review of Cooperative Learning Effects on High School and College Chemistry Achievement. *Journal of Chemical Education, 77*(1), 116–119

Abstract: This paper has two purposes. First, the reader is given an overview on how quantitative literature reviews (meta-analyses) can be conducted to give overall estimates of the quantitative impact an instructional treatment has on a specific student outcome. The second purpose is to illustrate how such a literature review is carried out by examining studies on using cooperative learning to teach chemistry at the high school and college levels. This analysis extends earlier reported work on effects of cooperative learning on achievement in college-level science, mathematics, and engineering and technology (SMET) courses. The analysis shows that while median student performance in a traditional course is at the 50th percentile, the median student performance in a cooperative learning environment is 14 percentile points higher.

Felder, R. M., Felder, G. N., and Dietz, E. J. (1998). A Longitudinal Study of Engineering Student Performance and Retention. V. *Comparisons with Traditionally-Taught Students. Journal of Engineering Education, 98*(4), 469–480

Abstract: In a longitudinal study at North Carolina State University, a cohort of students took five chemical engineering courses taught by the same instructor in five consecutive semesters. The courses made extensive use of active and cooperative learning and a variety of other techniques designed to address a broad spectrum of learning styles. Previous reports on the study summarized the instructional methods used in the experimental course sequence, described the performance of the cohort in the introductory chemical engineering course, and examined performance and attitude differences between students from rural and urban backgrounds and between male and female students [1–4]. This paper compares outcomes for the experimental cohort with outcomes for students in a traditionally-taught comparison group. The experimental group outperformed the comparison group on a number of measures, including retention and graduation in chemical engineering, and many more of the graduates in this group chose to pursue advanced study in the field. Since the experimental instructional model did not require small classes (the smallest of the experimental classes had 90 students) or specially equipped classrooms, it should be adaptable to any engineering curriculum at any institution.

Terenzini, P. T., Cabrera, A. F., Colbeck, C. L., Parente, J. M., and Bjorklund, S. A. (2001). Collaborative Learning vs. Lecture/Discussion: Students' Reported Learning Gains. *Journal of Engineering Education, 90*(1), 123–130

Abstract: This study examined the extent to which undergraduate engineering courses taught using active and collaborative learning methods differ from traditional lecture and discussion courses in their ability to promote the development of students' engineering design, problem-solving, communication, and group participation skills. Evidence for the study comes from 480 students enrolled in 17 active or collaborative learning courses/sections and six traditional courses/sections at six engineering schools. Results indicate that active or collaborative methods produce both statistically significant and substantially greater gains in student learning than those associated with more traditional instructional methods. These learning advantages remained even when differences in a variety of student pre-course characteristics were controlled.

Bonsangue, M. (1994). An efficacy study of the calculus workshop model. *CBMS Issues in Collegiate Mathematics Education, 4*, Providence, RI: American Mathematical Society, 117–137

Active Learning Pedagogies

Burrowes, P. A. (2003). A Student-Centered Approach to Teaching General Biology That Really Works: Lord's Constructivist Model Put to a Test. *The American Biology Teacher, 65*(7), 491–502

Abstract: None

Quote: Exams were taken on computer sheets and graded electronically. Grades achieved by students in experimental (constructivist teaching) and control (traditional teaching) groups were contrasted graphically (Figure 7) and the mean test scores were compared statistically by students' T-tests using Minitab 12 software. The first partial exam was offered after six weeks of instruction and included content on the cell as the functional unit of life (atoms, molecules, the cell membrane, organelles, energy transformations, cellular respiration, and photosynthesis). Although average scores of students in the experimental group were significantly better than in the control group ([Mean] = 65% versus 58%; T = 2.65, P = 0.004, n = 204), and they attained more As and Bs, and fewer Fs, the differences are not as impressive as later in the semester (Figure 7). The second exam was given 12 weeks into the semester, and evaluated knowledge on the continuity of life (mitosis, meiosis, DNA structure and replication, protein synthesis, and inheritance). Results of this exam showed grade improvement in both groups (Figure 7); however, the mean score of students in the experimental group was significantly higher than that of students in the control section ([Mean] = 72% versus 67%; T = 2.41, P = 0.009, n = 192). The outcome of the third exam (evolution and origin of life) was striking because performance of students in the experimental group approximated an ideal normal distribution of grades (Figure 7). Although students' achievement in the control group improved, students in the experimental section still did significantly better ([Mean] = 74% versus 68%; T = 3.05, P = 0.001 n =190).

Quote: This study provides substantiated evidence that teaching in a constructivist, active learning environment is more effective than traditional instruction in promoting academic achievement, increasing conceptual understanding, developing higher level thinking skills, and enhancing students interest in biology. In their final course evaluations, students in the experimental section commented that they enjoyed this class much more than their traditional classes, felt they had learned more, made valuable friendships in their collaborative groups and – particularly important to me – they never fell asleep! Thus, I am convinced that constructivism works better for our generation of students, and I will never return to a traditional style of teaching. Although the constructivist method of instruction requires a greater investment of time and effort from the professor for preparation, organization, and grading, the majority of this investment is made the first semester of teaching. During subsequent semesters, effort/payback increases dramatically, as less time is required. For example, with experience, I have become more efficient at formulating questions and coming up with ideas for problems, scenarios, and case studies, which help students develop their own knowledge of the material. Additionally, help from trained teaching assistants in grading, book-keeping, and organizational tasks associated with instruction can reduce some of the workload required of the instructor.

Laws, P., Sokoloff, D., and Thornton, R. (1999). Promoting Active Learning Using the Results of Physics Education Research. UniServe Science News, 13, Retrieved 4 September 2006 from http://science.uniserve.edu.au/newsletter/vol13/sokoloff.html

Redish, E. F., Saul, J. M., and Steinberg, R. N. (1997). On the effectiveness of active-engagement microcomputer-based laboratories. *American Journal of Physics, 65*(1), 45–54

Abstract: One hour active-engagement tutorials using microcomputer-based laboratory (MEL) equipment were substituted for traditional problem-solving recitations in introductory calculus-based mechanics classes for engineering students at the University of Maryland. The results of two specific tutorials, one on the concept of instantaneous velocity and one on Newton's third law were probed by using standard multiple-choice questions and a free-response final exam question. A comparison of the results of 11 lecture classes taught by six different teachers with and without tutorials shows that the MBL tutorials resulted in a significant improvement compared to the traditional recitations when measured by carefully designed multiple-choice problems. The free-response question showed that, although the tutorial students did somewhat better in recognizing and applying the concepts, there is still room for improvement.

Cummings, K., Marx, J., Thornton, R., and Kuhl, D. (1999). Evaluating innovations in studio physics. *American Journal of Physics, 67*(supplement 1 to no. 7), S38–S44

Abstract: In 1993, Rensselaer introduced the first Studio Physics course. Two years later, the Force Concept Inventory (FCI) was used to measure the conceptual learning gain [g] in the course. This was found to be a disappointing 0.22, indicating that Studio Physics was no more effective at teaching basic Newtonian concepts than a traditional course. Our study verified that result, [g(FCI,98)] = 0.18+/-0.12(s.d.), and thereby provides a baseline measurement of conceptual learning gains in Studio Physics I for engineers. These low gains are especially disturbing because the studio classroom appears to be interactive and instructors strive to incorporate modern pedagogies. The goal of our investigation was to determine if incorporation of research-based activities into Studio Physics would have a significant effect on conceptual learning gains. To measure gains, we utilized the Force Concept Inventory and the Force and Motion Conceptual Evaluation (FMCE). In the process of pursuing this goal, we verified the effectiveness of Interactive Lecture Demonstrations [[g(FCI)] =0.35+/-0.06(s.d.) and [g(FMCE)] = 0.45+/-0.03 (s.d.)] and Cooperative Group Problem Solving ([g(FCI)] = 0.36 and [g(FMCE)] = 0.36), and examined the feasibility of using these techniques in the studio classroom. Further, we have assessed conceptual learning in the standard Studio Physics course [[g(FCI,98)] = 0.18+/-0.12(s.d.) and [g(FMCE,98)] = 0.21+/-0.05 (s.d.)]. In this paper, we will clarify the issues noted above. We will also discuss difficulties in implementing these techniques for first time users and implications for the future directions of the Studio Physics courses at Rensselaer.

Hoellwarth, C., Moelter, M. J., and Knight, R. D. (2005). A direct comparison of conceptual learning and problem solving ability in traditional and studio style classrooms. *American Journal of Physics, 73*(5), 459–462

Abstract: We present data on student performance on conceptual understanding and on quantitative problem-solving ability in introductory mechanics in both studio and traditional classroom modes. The conceptual measures used were the Force Concept Inventory and the Force and Motion Conceptual Evaluation. Quantitative problem-solving ability was measured with standard questions on the final exam. Our data compare three different quarters over the course of 2 years. In all three quarters, the normalized learning gain in conceptual understanding was significantly larger for students in the studio sections. At the same time, students in the studio sections performed the same or slightly worse on quantitative final exam problems.

Beichner, R. J., Saul, J. M., Abbott, D. S., Morse, J. J., Deardorff, D. L., Allain, R. J., Bonham, J. W., Dancy, M. H., and Risley, J. S. (2007). The Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) Project. Retrieved August 27, 2007, from http://www.compadre.org/Repository/document/ServeFile.cfm?ID=4517&DocID=183

Abstract: The primary goal of the SCALE-UP Project is to establish a highly collaborative, hands-on, computer-rich, interactive learning environment for large, introductory college courses. North Carolina State University and a group of more than two-dozen collaborating schools are folding together lecture and lab with multiple instructors in a way that provides an effective, economical alternative to traditional lecture-oriented instruction. The project involves the development of the pedagogy, classroom environment, and teaching materials that will support this type of learning. It includes the development, evaluation, and dissemination of new curricular materials in physics, chemistry, and biology. Here we will focus on the calculus-based introductory physics part of the effort. In comparisons to traditional instruction we have seen significantly increased conceptual understanding, improved attitudes, successful problem solving, and higher success rates, particularly for females and minorities. This chapter highlights the development of the SCALE-UP pedagogy, classroom environment, and teaching materials for calculus-based introductory physics at North Carolina State University.

Summary: SCALE-UP pedagogy is characterized by the following common elements: (i) cooperative learning ("classroom renovated to emphasize group work with 2-3 groups of 3-4 students each per table"), (ii) active learning ("the majority of class time is spent on learning physics through activities done by groups of 3-4 students each", "the activities tend to be short (5-20 minutes) and followed by a class discussion"), (iii) research-based ("activities are based-on or at least informed by [physics education research] (PER)" (iv) classrooms that integrate "lecture and group work including experiments" with the following features: students work in groups of 2–4 students, access to computers and internet, access to equipment to perform experiments, facilitate class discussions, and share work among peers.

Quote: The SCALE-UP class demonstrated better improvement in conceptual understanding than Lecture/Laboratory classes by achieving higher normalized gains for the Mechanics semester pre/post force and motion concept tests at Coastal Carolina University (CCU), North Carolina State University (NCSU), University of Central Florida (UCF), University of New Hampshire (UNH), and Rochester Institute of Technology (RIT).

Quote: They report a 2-3x improvement in normalized gain on pre/post conceptual learning assessments such as the Force Concept Inventory, the Force and Motion Conceptual Evaluation, Conceptual Survey of Electricity and Magnetism, and the Electric Circuit Conceptual Evaluation [see Figures 5 and 6].

Quote: Failure rates are drastically reduced (typically 50%), especially for women and minorities [see Figure 7 and Table 6]

Michael, J. (2006). Where's the evidence that active learning works? *Advances in Physiology Education, 30*, 159–167

Abstract: Calls for reforms in the ways we teach science at all levels, and in all disciplines, are wide spread. The effectiveness of the changes being called for, employment of student-centered, active learning pedagogy, is now well supported by evidence. The relevant data have come from a number of different disciplines that include the learning sciences, cognitive

psychology, and educational psychology. There is a growing body of research within specific scientific teaching communities that supports and validates the new approaches to teaching that have been adopted. These data are reviewed, and their applicability to physiology education is discussed. Some of the inherent limitations of research about teaching and learning are also discussed.

Knight, J. K., and Wood, W. B. (2005). Teaching More by Lecturing Less. *Cell Biology Education, 4*, 298–310.

Abstract: We carried out an experiment to determine whether student learning gains in a large, traditionally taught, upper-division lecture course in developmental biology could be increased by partially changing to a more interactive classroom format. In two successive semesters, we presented the same course syllabus using different teaching styles: in fall 2003, the traditional lecture format; and in spring 2004, decreased lecturing and addition of student participation and cooperative problem solving during class time, including frequent in-class assessment of understanding. We used performance on pretests and posttests, and on homework problems to estimate and compare student learning gains between the two semesters. Our results indicated significantly higher learning gains and better conceptual understanding in the more interactive course. To assess reproducibility of these effects, we repeated the interactive course in spring 2005 with similar results. Our findings parallel results of similar teaching-style comparisons made in other disciplines. On the basis of this evidence, we propose a general model for teaching large biology courses that incorporates interactive engagement and cooperative work in place of some lecturing, while retaining course content by demanding greater student responsibility for learning outside of class.

Quote: The most compelling support for superiority of the interactive approach came from comparisons of normalized learning gains calculated from pretest and posttest scores in the traditional and interactive classes (Table 4). Normalized learning gain is defined as the actual gain divided by the possible gain, expressed as a percentage [100 × (posttest - pretest)/(100 - pretest); (Fagan et al., 2002)]. Normalization allows valid comparison and averaging of learning gains for students with different pretest scores. A comparison of the F’03 and S’04 courses showed a significant 16% difference ($p = .001$) in average learning gains, corresponding to a 33% improvement in performance by students in the more interactive S’04 course. Learning gains of greater than 60% were achieved by substantially more students in the interactive class (43/70) than in the traditional class (19/72) (Figures 2, 3). Both ‘‘A’’ and ‘‘B’’ students made higher gains in the interactive course, while ‘‘C’’ students achieved about the same learning gain range in both semesters (Figure 2).

Freeman, S., O’Connor, E., Parks, J. W., Cunningham, M., Hurley, D., Haak, D., Dirks, C., and Wenderoth, M. P., (2007). Prescribed Active Learning Increases Performance in Introductory Biology. *Cell Biology Education, 6*, 132–139.

Abstract: We tested five course designs that varied in the structure of daily and weekly active-learning exercises in an attempt to lower the traditionally high failure rate in a gateway course for biology majors. Students were given daily multiple-choice questions and answered with electronic response devices (clickers) or cards. Card responses were ungraded; clicker responses were graded for right/wrong answers or participation. Weekly practice exams were done as an individual or as part of a study group. Compared with previous versions of the same course taught by the same instructor, students in the new course designs performed better: There were significantly lower failure rates, higher total exam points, and higher scores on an identical midterm. Attendance was higher in the clicker versus cards section; attendance and course grade were positively correlated. Students did better on clicker questions if they were graded for right/wrong answers versus participation, although this improvement did not translate into increased scores on exams. In this course, achievement increases when students get regular practice via prescribed (graded) active-learning exercises.

Problem-based Learning Pedagogies

Prince, M. J., and Felder, R. M. (2006). Inductive Teaching and Learning Methods: Definitions, Comparisons, and Research Bases. *Journal of Engineering Education, 95*(2), 123–138.

Abstract: Traditional engineering instruction is deductive, beginning with theories and progressing to the applications of those theories. Alternative teaching approaches are more inductive. Topics are introduced by presenting specific observations, case studies or problems, and theories are taught or the students are helped to discover them only after the need to know them has been established. This study reviews several of the most commonly used inductive teaching methods, including inquiry learning, problem-based learning, project-based learning, case-based teaching, discovery learning, and just-in-time teaching. The paper defines each method, highlights commonalities and specific differences, and reviews research on the effectiveness of the methods. While the strength of the evidence varies from one method to another, inductive methods are consistently found to be at least equal to, and in general more effective than, traditional deductive methods for achieving a broad range of learning outcomes.

Quote: Individual studies have found a robust positive effect of PBL [problem-based learning] on skill development [1, 65, 66], understanding the interconnections among concepts [65], deep conceptual understanding [67], ability to apply appropriate metacognitive and reasoning strategies [68], teamwork skills [69], and even class attendance [70], but have not reached any firm conclusion about the effect on content knowledge. A longitudinal study of the effectiveness of the McMaster PBL program in chemical engineering demonstrated its superiority to traditional education in the development of key process skills [55]. PBL has also been shown to promote self-directed learning [71] and the adoption of a deep (meaning-oriented) approach to learning, as opposed to a superficial (memorization-based) approach [21, 46, 72].

Prince, M., and Felder, R. (2007). The Many Faces of Inductive Teaching and Learning. *Journal of College Science Teaching, 36*(5), 14–20

Tag: This study examines the effectiveness and implementation of different inductive teaching methods, including inquiry-based learning, problem-based learning, project-based learning, case-based teaching, discovery learning, and just-in-time teaching.

Dochy, F., Segers M., Van den Bossche, P., and Gijbels, D. (2003). Effects of Problem-Based Learning: A Meta-Analysis. *Learning and Instruction, 13*, 533–568

Abstract: This meta-analysis has two aims: (a) to address the main effects of problem based learning on two categories of outcomes: knowledge and skills; and (b) to address potential moderators of the effect of problem based learning. We selected 43 articles that met the criteria for inclusion: empirical studies on problem based learning in tertiary education conducted in real-life classrooms. The review reveals that there is a robust positive effect from PBL on the skills of students. This is shown by the vote count, as well as by the combined effect size. Also no single study reported negative effects. A tendency to negative results is discerned when considering the effect of PBL on the knowledge of students. The combined effect size is significantly negative. However, this result is strongly influenced by two studies and the vote count does not reach a significant level. It is concluded that the combined effect size for the effect on knowledge is non-robust. As possible moderators of PBL effects, methodological factors, expertise-level of students, retention period and type of assessment method were investigated. This moderator analysis shows that both for knowledge- and skills-related outcomes the expertise-level of the student is associated with the variation in effect sizes. Nevertheless, the results for skills give a consistent positive picture. For knowledge-related outcomes the results suggest that the differences encountered in the first and the second year disappear later on. A last remarkable finding related to the retention period is that students in PBL gained slightly less knowledge, but remember more of the acquired knowledge.

Quote: For skill development, the results are unequivocal: 14 studies found a positive effect and none found a negative effect, and the weighted average effect size was 0.460(±0.058).

Quote: For knowledge acquisition, seven of the studies analyzed found a positive effect and 15 found a negative effect, with weighted average effect size and 95 percent confidence interval -0.223 (±0.058). When the assessment of knowledge is carried out some time after the instruction was given, the effect of PBL positive.

Gijbels, D., Dochy, F., Van den Bossche, P., and Segers, M. (2005). Effects of Problem-Based Learning: A Meta-Analysis from the Angle of Assessment. *Review of Educational Research, 75*(1), 27–61

Abstract: This meta-analysis investigated the influence of assessment on the reported effects of problem-based learning (PBL) by applying Sugrue's (1995) model of cognitive components of problem solving. Three levels of the knowledge structure that can be targeted by assessment of problem solving are used as the main independent variables: (a) understanding of concepts, (b) understanding of the principles that link concepts, and (c) linking of concepts and principles to conditions and procedures for application. PBL had the most positive effects when the focal constructs being assessed were at the level of understanding principles that link concepts. The results suggest that the implications of assessment must be considered in examining the effects of problem-based learning and probably in all comparative education research.

Quote: Three levels of the knowledge structure in assessment of problem solving: (a) understanding of concepts; (b) understanding of the principles that link concepts; (c) linking of concepts and principles to conditions and procedures for application. PBL had the most positive effects when the focal constructs being assessed were at the level of understanding principles that link concepts.

Vernon, D. T. A., and Blake, R. L. (1993). Does Problem-Based Learning Work? A Meta-Analysis of Evaluative Research. *Academic Medicine, 68*, 550–563.

Abstract: The purpose of this review is to synthesize all available evaluative research from 1970 through 1992 that compares problem-based learning (PBL) with more traditional methods of medical education. Five separate meta-analyses were performed on 35 studies representing 19 institutions. For 22 of the studies (representing 14 institutions), both effect-size and supplementary vote-count analyses could be performed; otherwise, only supplementary analyses were performed. PBL was found to be significantly superior with respect to students' program evaluations (i.e., students' attitudes and opinions about their programs)--dw (standardized differences between means, weighted by sample size) = +.55, CI.95 = +.40 to +.70 - and measures of students' clinical performance (dw = +.28, CI.95 = +.16 to +.40). PBL and traditional methods did not differ on miscellaneous tests of factual knowledge (dw = -.09, CI.95 = +.06 to -.24) and tests of clinical knowledge (dw = +.08, CI.95 = -.05 to +.21). Traditional students performed significantly better than their PBL counterparts on the National Board of Medical Examiners Part I examination--NBME I (dw = -.18, CI.95 = -.10 to -.26). However, the NBME I data displayed significant overall heterogeneity (Qt = 192.23, $p < .001$) and significant differences among programs (Qb = 59.09, $p < .001$), which casts doubt on the generality of the findings across programs. The comparative value of PBL is also supported by data on outcomes that have been studied less frequently, i.e., faculty attitudes, student mood, class attendance, academic process variables, and measures of humanism. In conclusion, the results generally support the superiority of the PBL approach over more traditional methods.

Capon, N., and Kuhn, D. (2004). What's So Good About Problem-Based Learning? *Cognition and Instruction, 22*(1), 61–79

Abstract: In a systematically designed and controlled experiment conducted in a naturalistic instructional setting, we examined adult students' learning of two concepts. Two intact classes taught by the same instructor were assigned to 1 of 2 conditions. In 1 class, instruction was problem based for 1 concept. For a second concept, lecture/discussion was the exclusive method. In the other class, matching of concept and method (problem based or lecture/discussion) was reversed. Two forms of assessment of learning occurred 6 and 12 weeks following instruction. At the initial assessment, the lecture/discussion group showed superior learning for 1 concept and the groups performed equivalently for the other concept. At the later assessment, however, the 2 groups showed equivalent ability to access each of the concepts, but each group showed superior explanation of the concept for which they had experienced problem-based learning. Results support the hypothesis of integration of new information with existing knowledge structures activated by the problem-based experience as the mechanism by which problem-based learning produces its benefits.

Inquiry-based Learning Pedagogies

Prince, M. J., and Felder, R. M. (2006). Inductive Teaching and Learning Methods: Definitions, Comparisons, and Research Bases. *Journal of Engineering Education, 95*(2), 123–138.

Abstract: Traditional engineering instruction is deductive, beginning with theories and progressing to the applications of those theories. Alternative teaching approaches are more inductive. Topics are introduced by presenting specific observations, case studies or problems, and theories are taught or the students are helped to discover them only after the need to know them has been established. This study reviews several of the most commonly used inductive teaching methods, including inquiry learning, problem-based learning, project-based learning, case-based teaching, discovery learning, and just-in-time teaching. The paper defines each method, highlights commonalities and specific differences, and reviews research on the effectiveness of the methods. While the strength of the evidence varies from one method to another, inductive methods are consistently found to be at least equal to, and in general more effective than, traditional deductive methods for achieving a broad range of learning outcomes.

Farrell, J. J., Moog, R. S., and Spencer, J. N. (1999). A Guided Inquiry General Chemistry Course. *Journal of Chemical Education, 74*(4), 570–574

Abstract: A first-year general chemistry course based on constructivist principles and the learning cycle has been developed. Through the use of cooperative learning techniques, students are active participants in the learning process. No lectures are given; students follow guided inquiry worksheets to develop and understand the course concepts. Groups of about four students are formed and the instructor moves among the groups, serving as a facilitator. The laboratory is designed in the same way as the classroom component of the course. The students form hypotheses and collect data, leading to further refinement of the hypotheses and to formation of chemical concepts.

Quote: the sections taught according to the principles of guided inquiry have experienced a decrease in the W, D, F rate from 21.9% (420 students, Fall 1990–Spring 1994) to 9.6% (438 students, Fall 1994–Fall 1997).3 In the Guided Inquiry (GI) sections, the withdrawal rate is 2.3% and only 1 out of 438 students has received a grade of F in these sections. In contrast, students taught by these same instructors previously had a W rate of 9.3% and 3.6% failed. Final exams given to the GI students that were substantially similar to exams given in the past showed that GI students scored as high as or higher than students who had taken a more traditional course from the same instructor.

Lewis, S. E., and Lewis, J. E. (2005). Departing from Lectures: An Evaluation of a Peer-Led Guided Inquiry Alternative. *Journal of Chemical Education, 82*(1), 135–139

Abstract: To improve a large-enrollment general chemistry course based on conventional lectures, we instituted a reform combining peer-led team learning with a guided inquiry approach, together called peer-led guided inquiry (PLGI). For one group of first-semester general chemistry students, a PLGI session was combined with two lectures per week, and this group was compared to a control group that had the usual three lectures per week. Students were compared based on performance on identical course exams and on a final exam from the ACS Examinations Institute given at the end of the semester. The experimental group was found to

perform better than the control group overall, in spite of experiencing one fewer lecture each week. Also, attendance at the PLGI sessions was found to have a significant positive impact on student performance, even when controlling for students' SAT mathematics and verbal scores. This method of evaluating reform effects for institutions with several large sections of introductory chemistry courses is recommended.

Quote: the experimental group consistently outperformed the control group on the course exams and on the final exam

Quote: The results from this analysis indicate that a student who attends PLGI [peer-led guided inquiry] sessions can be expected to perform better on exams than another student at the same SAT level.

Other results for the Process Oriented Guided Inquiry Learning (POGIL) project are available at http://www.pogil.org/effectiveness/.

Challenge-based Learning Pedagogies

Roselli, R. J., and Brophy, S. P. (2006). Effectiveness of Challenge-Based Instruction in Biomechanics. *Journal of Engineering Education, 95*(4), 311–324.

Abstract: Studies were designed to determine the effectiveness of challenge-based instruction (CBI) versus traditional lecture-based instruction. Comparisons were made over a three-year period between student performance on knowledge-based questions in courses taught with taxonomy-based and challenge-based approaches to instruction. When performance on all questions was compared, CBI classes scored significantly better than control classes on 26 percent of the questions, while control classes outperformed CBI classes on eight percent of the questions, but there was no significant difference in overall performance. However, students in CBI classes performed significantly better than students in control classes on the more difficult questions (35 percent versus four percent). We attribute these differences to additional opportunities available in CBI classrooms for learners to examine their conceptual understanding. Student surveys indicate a slight preference for the challenge-based approach. We believe that the challenge-based approach is effective and has the potential to better prepare students for the workplace and for life-long learning.

Peer-Led Team Learning Pedagogies

Tien, L. T., Roth, V., and Kampmeier, J. A. (2001). Implementation of a Peer-Led Team Learning Instructional Approach in an Undergraduate Organic Chemistry Course. *Journal of Research in Science Teaching, 39*(7), 606–632

Abstract: This study focuses on the implementation of a peer-led team learning (PLTL) instructional approach for all students in an undergraduate organic chemistry course and the evaluation of student outcomes over 8 years. Students who experienced the student-centered instruction and worked in small groups facilitated by a peer leader (treatment) in 1996–1999 were compared with students who experienced the traditional recitation section (control) in 1992–1994. Quantitative and qualitative data show statistically significant improvements in student performance, retention, and attitudes about the course. These findings suggest that using

undergraduate leaders to implement a peer-led team learning model that is built on a social constructivist foundation is a workable mechanism for effecting change in undergraduate science courses.

McCreary, C. L., Golde, M. F., and Koeske, R. (2006). Peer Instruction in the General Chemistry Laboratory: Assessment of Student Learning. *Journal of Chemical Education, 83*(5), 804–810

Abstract: We report the first systematic comparison of conventional and workshop labs. A natural experiment proved possible because students sign up for labs without knowing the type of instruction they will receive. A reliable grading system was developed to characterize students' written responses to the final lab exam, and an independent rater used it to assess student learning. Assessments of learning were made without knowledge of students' instructional condition. Compared to students in conventional sections, students in workshop sections showed superior learning and critical thinking skills, and gave answers that were longer and of greater clarity. Possible reasons for these improvements are discussed.

Quote: In fact, for nearly every measure of performance quality and written communication included in the study, participation in the Workshop labs tended to enhance students' learning relative to that indicated by test performance for students in the conventionally taught labs, with the differences often reaching statistical significance.

Other results from the Peer-Led Team Learning Workshop Project are available at http://www.sci.ccny.cuny.edu/~chemwksp/ResearchAndEvaluationComparisons.html

Workshop Groups

Born, W. K., Revelle, W., and Pinto, L. H. (2002). Improving Biology Performance with Workshop Groups. *Journal of Science Education and Technology, 11*(4), 347– 365

Abstract: This 2-year quasi-experiment evaluated the effect of peer-led workshop groups on performance of minority and majority undergraduate biology students. The workshop intervention used was modeled after a program pioneered by Treisman (1992). Majority volunteers randomly assigned to workshops ($n = 61$) performed significantly better than those assigned to the control group ($n = 60$, $p < 0.05$) without spending more time studying. Workshop minority students ($n = 25$) showed a pattern of increasing exam performance in comparison to historic control minority students ($n = 21$), who showed a decreasing pattern ($p < 0:05$). Volunteers ($n = 121$) initially reported that biology was more interesting and more important to their futures than to nonvolunteers' ($n = 435$, $p < 0.05$). Volunteers also reported higher levels of anxiety related to class performance ($p < 0.05$). The relationship of anxiety to performance was moderated by volunteer status. Performance of volunteers was negatively associated with self-reported anxiety ($r = –0.41$, $p < 0.01$). Performance of nonvolunteers was unrelated to self-reported anxiety ($r = –0.02$). Results suggest elevated anxiety related to class performance may increase willingness to participate in activities such as workshop interventions. In addition, students who volunteer for interventions such as workshops may be at increased risk of performance decrements associated with anxiety. Even so, workshop programs appear to be an effective way to promote excellence among both majority and minority students who volunteer

to participate, despite the increased risk of underperformance associated with higher levels of anxiety.

Classroom Assessment Techniques

Chizmar, J. F., and Ostrosky, A. L. (1998). The One-Minute Paper: Some Empirical Findings.: *The Journal of Economic Education, 29*(1), 3–10

Abstract: A major finding of the Harvard Assessment Seminars is that "modest, relatively simple and low-tech innovations can improve students' learning and active participation in class" (Light 1990, 6). One such innovation is the so-called one-minute paper (Cross and Angelo 1988; Bateman and Roberts 1992a, 1992b; Wilson 1986). The one-minute paper has become rather ubiquitous in higher education. According to Cross and Angelo (1993, 148), "No other Classroom Assessment Technique[1] has been used more often or by more college teachers than the [One] Minute Paper." When asked by college teachers to identify the single pedagogical innovation that would most improve their teaching, Light (1990, 35) always responds with the one-minute paper, an idea that "swamped all others." In this article, we describe the one-minute paper and report on a pilot implementation of this technique to manage instruction in the micro portion of the introductory economics course at a large public university. We conclude with a discussion of issues and questions revealed by the pilot implementation that may affect the efficacy of the one-minute paper.

Findings: "This result suggested, as we hypothesized, that the use of the one-minute paper improves student performance. Its coefficient implied that the use of the one-minute paper increased student performance by approximately .5 of a point on the postTUCE exam, ceteris paribus."

Findings: "This evidence suggests that the benefit to students from using the one-minute paper does not depend on the instructor who implements it."

Findings: "This evidence supported our initial hypothesis that the benefit to students from using the one-minute paper does not depend on their ability level."

Almer, E. D., Jones, K., and Moeckel, C. L. (1998). The impact of one-minute papers on learning in an introductory accounting course. *Issues in Accounting Education, 13*(3), 485–495

Abstract: This study examines the potential performance benefits of an often-cited pedagogical tool: one-minute papers. The effect of various forms of one-minute papers on quiz scores was investigated in an undergraduate introductory account course with over 850 students. Students were required to write one-minute papers addressing (1) the main point learned in class and (2) the main unanswered question from class that day. Overall results indicate that performance on subsequent essay quizzes was significantly higher by students who wrote one-minute papers than performance by students who did not write the papers. Of particular interest to instructors was that the increase in quiz scores when one-minute papers were not graded was significantly higher than when the one-minute papers were graded. Results of this study should be useful to instructors interested in an efficient and effective pedagogical tool.

Formative Assessment

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7–74

Abstract: This article is a review of the literature on classroom formative assessment. Several studies show firm evidence that innovations designed to strengthen the frequent feedback that students receive about their learning yield substantial learning gains. The perceptions of students and their role in self-assessment are considered alongside analysis of the strategies used by teachers and the formative strategies incorporated in such systemic approaches as mastery learning. There follows a more detailed and theoretical analysis of the nature of feedback, which provides a basis for a discussion of the development of theoretical models for formative assessment and of the prospects for the improvement of practice.

Undergraduate Research Experiences

Seymour, E., Hunter, A.-B., Laursen, S. L., and Diatonic, T. (2004). Establishing the Benefits of Research Experiences for Undergraduates in the Sciences: First Findings from a Three-Year Study. *Science Education, 88*, 493–534

Abstract: Descriptions of student-identified benefits of undergraduate research experiences are drawn from analysis of 76 first-round student interviews gathered at the end of summer 2000 at four participating liberal arts colleges (Grinnell, Harvey Mudd, Hope, and Wellesley). As part of the interview protocol, students commented on a checklist of possible benefits derived from the literature. They also added gains that were not on this list. Students were overwhelmingly positive: 91% of all statements referenced gains from their experiences. Few negative, ambivalent, or qualified assessments of their research experiences were offered. The benefits described were of seven different kinds. Expressed as percentages of all reported gains, they were personal/professional gains (28%); "thinking and working like a scientist" (28%); gains in various skills (19%); clarification/confirmation of career plans (including graduate school) (12%); enhanced career/graduate school preparation (9%); shifts in attitudes to learning and working as a researcher (4%); and other benefits (1%).

Lopatto, D. (2004). Survey of Undergraduate Research Experiences (SURE): First Findings. *Cell Biology Education, 3*, 270–277.

Abstract: In this study, I examined the hypothesis that undergraduate research enhances the educational experience of science undergraduates, attracts and retains talented students to careers in science, and acts as a pathway for minority students into science careers. Undergraduates from 41 institutions participated in an online survey on the benefits of undergraduate research experiences. Participants indicated gains on 20 potential benefits and reported on career plans. Over 83% of 1,135 participants began or continued to plan for postgraduate education in the sciences. A group of 51 students who discontinued their plans for postgraduate science education reported significantly lower gains than continuing students. Women and men reported similar levels of benefits and similar patterns of career plans. Ethnic

groups did not significantly differ in reported levels of benefits or plans to continue with postgraduate education.

Hunter, A.-B., Laursen, S. L., and Seymour, E. (2007). Becoming a Scientist: The Role of Undergraduate Research in Students' Cognitive, Personal, and Professional Development. *Science Education, 91*, 36–74

Abstract: In this ethnographic study of summer undergraduate research (UR) experiences at four liberal arts colleges, where faculty and students work collaboratively on a project of mutual interest in an apprenticeship of authentic science research work, analysis of the accounts of faculty and student participants yields comparative insights into the structural elements of this form of UR program and its benefits for students. Comparison of the perspectives of faculty and their students revealed considerable agreement on the nature, range, and extent of students' UR gains. Specific student gains relating to the process of "becoming a scientist" were described and illustrated by both groups. Faculty framed these gains as part of professional socialization into the sciences. In contrast, students emphasized their personal and intellectual development, with little awareness of their socialization into professional practice. Viewing study findings through the lens of social constructivist learning theories demonstrates that the characteristics of these UR programs, how faculty practice UR in these colleges, and students' outcomes—including cognitive and personal growth and the development of a professional identity—strongly exemplify many facets of these theories, particularly, student-centered and situated learning as part of cognitive apprenticeship in a community of practice.

Russell, S. H., Hancock, M. P., and McCullough, J. (2007). Benefits of Undergraduate Research Experiences. *Science, 316*, 548–549

Quote: "We found that UROs [undergraduate research opportunities] increase understanding, confidence, and awareness (5–8). Most (88%) of the respondents to the NSF follow-up survey reported that their understanding of how to conduct a research project increased a fair amount or a great deal, 83% said their confidence in their research skills increased, and 73% said their awareness of what graduate school is like increased.

Lack of Growth in Valued Learning Outcomes in Traditional Curricula

Litzinger, T., Wise, J., Lee, S., Bjorklund, S. (2003). Assessing Readiness for Self-directed Learning, *Proceedings, ASEE Annual Conference & Exposition*, Retrieved from http://www.asee.org/acPapers/2003-1429_Final.pdf, 2 July 2003

Abstract: None

Quote: Figure 1 presents the average SDLRS scores for the five groups of students in the study, who were grouped by semester standing according to academic year from first year (1&2) to "supersenior" year (9&10). The average scores range from 217 to 228, corresponding to percentile ranks, based on SDLRS results for adults, of 50% and 68%, respectively. Although the data suggest a slight upward trend, the trend proved not to be statistically significant based upon an analysis of variance (ANOVA). Thus the cross-sectional study did not find evidence of an increase in readiness for self-directed learning, even for students in the later semesters who are taking elective courses and their capstone courses.

Woods, D. et al. (1997). Developing Problem Solving Skills: The McMaster Problem Solving Program. *Journal of Engineering Education, 86*(2), 75–92

Abstract: This paper describes a 25-year project in which we defined problem solving, identified effective methods for developing students' skill in problem solving, implemented a series of four required courses to develop the skill, and evaluated the effectiveness of the program. Four research projects are summarized in which we identified which teaching methods failed to develop problem solving skill and which methods were successful in developing the skills. We found that students need both comprehension of Chemical Engineering and what we call general problem solving skill to solve problems successfully. We identified 37 general problem solving skills. We use 120 hours of workshops spread over four required courses to develop the skills. Each skill is built (using content-independent activities), bridged (to apply the skill in the content-specific domain of Chemical Engineering) and extended (to use the skill in other contexts and contents and in everyday life). The tests and examinations of process skills, TEPS, that assess the degree to which the students can apply the skills are described. We illustrate how self-assessment was used.

Quote: During the four-year undergraduate engineering program studied, 1974-1978, the students had worked over 3000 homework problems, they had observed about 1000 sample solutions being worked on the board by either the teacher or by peers, and they had worked many open-ended problems.36 In other words, they showed no improvement in problem solving skills despite the best intentions of their instructors. Caillot[37] and Meiring[26] confirm these findings.

Fowler, D., Maxwell, D., and Froyd, J. (2003). Learning Strategy Growth Not What Expected After Two Years through Engineering Curriculum. *Proceedings, ASEE Annual Conference & Exposition*, Retrieved from http://www.asee.org/acPapers/2003-534_Final.pdf, 2 July 2003

Abstract: As the pace of technological development continues to increase, consensus has emerged that undergraduate science, technology, engineering and mathematics (STEM) curricula cannot contain all of the topics that engineering professionals will require, even during the first ten years of their careers. Therefore, the need for students to increase their capability for lifelong learning is receiving greater attention. It is anticipated that development of this capability occurs during the undergraduate curricula. However, preliminary data from both first-year and junior engineering majors may indicate that development of these competencies may not be as large as desired. Data was obtained using the Learning and Study Skills Inventory (LASSI), an instrument whose reliability has been demonstrated during the past fifteen years. The LASSI is a ten-scale, eighty-item assessment of students' awareness about and use of learning and study strategies related to skill, will and self-regulation components of strategic learning. Students at Texas A&M University in both a first-year engineering course and a junior level civil engineering course took the LASSI at the beginning of the academic year. Improvements would normally be expected after two years in a challenging engineering curriculum. However, data on several different scales appears to indicate that improvements are smaller than might be expected.

CBE—Life Sciences Education
Vol. 9, 10–16, Spring 2010

Feature
From the National Academies

Integrated Biology and Undergraduate Science Education: A New Biology Education for the Twenty-First Century?

Jay B. Labov,* Ann H. Reid,†‡ and Keith R. Yamamoto§

*Center for Education and †Board on Life Sciences, National Research Council, Washington, DC 20001; and §Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94143

INTRODUCTION

> *Given the radical changes in the nature of the science of biology and what we have learned about effective ways to teach, this is an opportune time to address the biology we teach so that it better represents the biology we do.*
>
> – www.visionandchange.org

For more than a decade, numerous reports have called for a rethinking and restructuring of high school and undergraduate science education to make it more relevant and accessible to a broader spectrum of students (Handelsman *et al*., 2006; Hulleman and Harackiewicz, 2009; National Research Council [NRC], 1996, 1997, 1998, 2002, 2003a,b,c, 2005, 2008; National Science Foundation [NSF], 1996) and to base our strategies on the expanding body of research on human learning and cognition (NRC, 2000b; Allen and Tanner, 2007; Morse and Jutras, 2008; DeHaan, 2009, Pfund *et al*., 2009, Labov *et al*., 2009). In 2009, several important publications, conferences, and events have pointed toward confluence around more interdisciplinary and interconnected approaches and themes for undergraduate education in the life sciences. These events have included the following:

- Release of draft curriculum frameworks in biology for the College Board's multiyear restructuring of advanced placement courses in science for high school students (see http://apcentral.collegeboard.com/apc/public/repository/draft_revised_ap_biology_curriculum.pdf). This restructuring closely follows the recommendations of a report from the NRC (2002) and calls for teaching fewer concepts in greater depth. Restructuring also requires developing and implementing means to measure students' level of conceptual understanding (Mervis, 2009a; Wood, 2009).
- Publication of *Scientific Foundations for Future Physicians*, a joint report from the Howard Hughes Medical Institute (HHMI) and the Association of American Medical Colleges, which calls for a change in undergraduate science education away from a system based on courses to one based on "competencies." According to the committee, "A competency-based approach will give both learners and educators more flexibility in the premedical curriculum and allow the development of more interdisciplinary and integrative courses that maintain scientific rigor, while providing a broad education." (Executive Summary, p. 1)[1]
- Convening of "Vision and Change in Undergraduate Biology Education," a summit held in Washington, DC, in July 2009 that was organized by the American Association for the Advancement of Science with support from the NSF. This summit brought together >500 people to consider future pathways for undergraduate education in the life sciences (Mervis, 2009b; Woodin *et al*., 2009).[2] A report from the summit is planned for release in 2010.
- Publication in September 2009 of *A New Biology for the Twenty-First Century* by a committee under the aegis of the NRC's Board on Life Sciences (NRC, 2009; a podcast about the report is available at http://dels.nas.edu/dels/viewreport.cgi?id=5953). The report proposes a bold new integrated research agenda, with important implications for the future of undergraduate and K–12 science education.
- Convening in November 2009 of an interdisciplinary forum on synthetic biology as part of the annual National Academies Keck Futures Initiative.[3] Consistent with calls to find ways to develop science curricula in conjunction with cutting-edge scientific discoveries (Jurkowski *et al*., 2007), the forum actively considered issues of education and communication about synthetic biology in conjunc-

DOI: 10.1187/cbe.09–12–0092
‡ Present address: American Academy of Microbiology, Washington, DC.
Address correspondence to: Jay B. Labov (jlabov@nas.edu).

[1] A separate Executive Summary for this report is available at www.hhmi.org/grants/pdf/08-209_exec_summary.pdf.
[2] Additional information is available at www.visionandchange.org.
[3] Additional information is available at www.keckfutures.org/conferences/synthetic-biology.html.

tion with discussions of scientific, legal, and ethical aspects. A report from this event will be published by the National Academies in 2010.

Thus, throughout this past year, the life sciences community has focused its attention on where biological research is likely to progress over the next several decades and how education in the life sciences might keep pace with this rethinking of research priorities and progress. The NRC (2009) report offers the most comprehensive review of these sets of issues; its recommendations for research and education agendas are summarized below.

A NEW BIOLOGY FOR THE TWENTY-FIRST CENTURY: OVERVIEW AND IMPLICATIONS FOR BIOLOGICAL RESEARCH

> *Biological research is in the midst of a revolutionary change due to the integration of powerful technologies along with new concepts and methods derived from inclusion of physical sciences, mathematics, computational sciences, and engineering. As never before, advances in biological sciences hold tremendous promise for surmounting many of the major challenges confronting the United States and the world. Historically, major advances in science have provided solutions to economic and social challenges. At the same time, those challenges have inspired science to focus its attention on critical needs. Scientific efforts based on meeting societal needs have laid the foundation for countless new products, industries, even entire economic sectors that were unimagined when the work began . . .*
>
> *. . . the essence of the New Biology is integration–reintegration of the many subdisciplines of biology, and the integration into biology of physicists, chemists, computer scientists, engineers, and mathematicians to create a research community with the capacity to tackle a broad range of scientific and societal problems. NRC (2009), p. viii*
>
> *. . . the New Biology represents an additional, complementary approach to biological research. Purposefully organized around problem-solving, this approach marshals the basic research to advance fundamental understanding, brings together researchers with different expertise, develops the technologies required for the task and coordinates efforts to ensure that gaps are filled, problems solved, and resources brought to bear at the right time.*
>
> – NRC, 2009, p. 3

The committee[4] that authored *A New Biology* (NRC, 2009; Figure 1) was asked by the National Institutes of Health, NSF, and the U.S. Department of Energy to undertake an appraisal of areas in which the life sciences are poised to make major advances and of how these advances could contribute to practical applications and improved environmental stewardship, human health, and quality of life. It also was asked to examine current trends toward integration and synthesis within the life sciences, the increasingly important role of interdisciplinary teams, and the resultant implications for funding strategies, decision making, infrastructure, and education in the life sciences.

[4] A list of committee members and their institutional affiliations is available at http://books.nap.edu/openbook.php?record_id=12764&page=R5.

The report states that the life sciences face a moment of opportunity similar to that faced by physics in the twentieth century. The members of the committee identified four major areas of societal challenge where problem-focused research incorporating emerging theory, new technologies, fundamental findings from basic research in the life sciences, and integration into the life sciences of the physical sciences, mathematics, and engineering could enable biology to contribute to rapid progress in practical problem-solving. These broad areas, which are in fact interdependent and must be addressed in parallel, include the following:

- health, with an emphasis on developing the capacity to understand individual health at a level that allows prevention, diagnosis, and treatment to be based on each individual's unique genetic and environmental characteristics rather than statistical probability;
- environment, with an emphasis on developing the means to monitor, diagnose, and restore ecosystem function and biodiversity in the face of rapid environmental change;
- energy, with an emphasis on expanding sustainable alternatives to fossil fuels; and
- food, with an emphasis on developing the capability to adapt any crop plant to sustainable growth under any set of growing conditions. The new biology, if successful, would make it possible to more quickly and predictably breed food plants suitable for cultivation where they are most needed.

The committee envisioned the New Biology as a cycle encompassing four major components (Figure 2):

1. Integration of Scientific Information, Theory, Technologies, and Thinking about Complex Problems. As noted in Figure 2, biology is essential, but in its traditional form is insufficient to confront the key problems that must be addressed in the future. The physical sciences, mathematics, engineering, and information sciences all must be integrated with the traditional discipline to form the New Biology. Importantly, the committee emphasized that science education must be an integral input to this interdisciplinary approach to capacious problems. Science education itself also is envisioned as advancing as a result of the feedback loops that emerge from this integrated approach.
2. Deeper Understanding of Biological Systems. A deeper understanding of biological systems emerges from the multifaceted thinking of experts from a variety of disciplines. This deeper understanding will advance biology from an era of observation and mechanism to one of deciphering design principles for biological processes, making them accessible to manipulation and eventually predictable.
3. Biologically Based Solutions to Societal Problems. For societal problems that may be intractable by other approaches, the deeper understanding that results from the integrated and interdisciplinary collaborations driving the New Biology will allow more rapid progress on complex and interrelated challenges such as those in the areas of health, environment, energy, and food. In this context, the societal issues could be

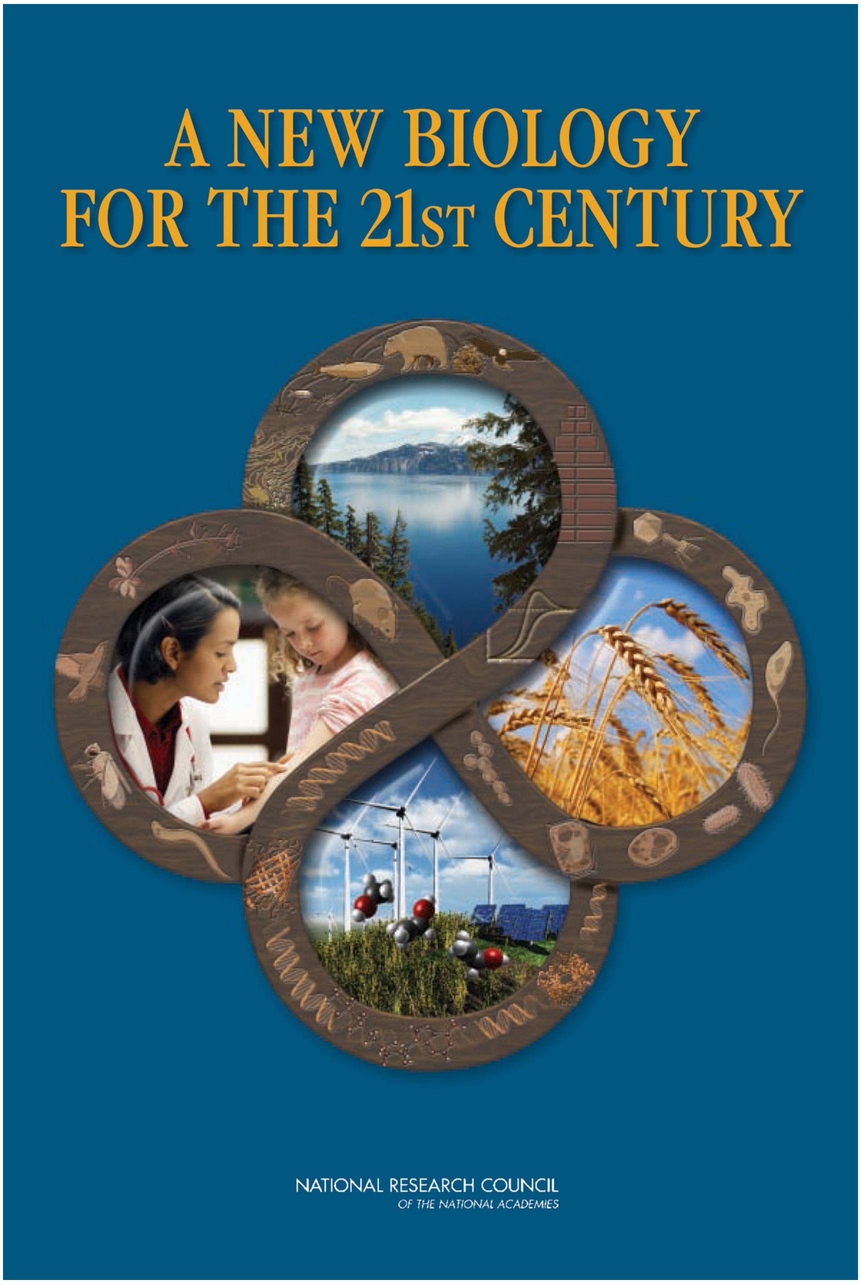


Figure 1. Report cover for *A New Biology for the 21st Century*.

considered as interactive drivers on a very large scale, spurring the development of enabling technologies and new discovery.

4. Feedback and Benefits to Contributing Disciplines and to Education. The collective, synergistic knowledge and thinking that emerge from integrated approaches to biological research and their applications to societal challenges will, in turn, inform and stimulate fundamental research across the scientific spectrum and in science education. If education tracks the projected trajectory of research that is encompassed by the New Biology, individual disciplines are also likely to converge around the idea of integrated and interconnected science, technology, engineering, and mathematics (STEM) education.

A NEW BIOLOGY FOR THE TWENTY-FIRST CENTURY: OVERVIEW AND IMPLICATIONS FOR BIOLOGICAL EDUCATION

The committee observed that the New Biology presents unprecedented opportunities to draw attention to the excitement of biology but will require new ways of thinking about

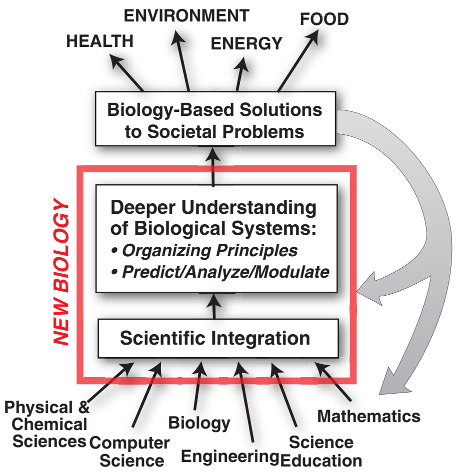


Figure 2. The inputs to and outcomes of a new integrated approach to biological research in the twenty-first century (NRC, 2009, p. 18).

how to attract, educate, and retain undergraduates as detailed below.

The New Biology Initiative Provides an Opportunity to Attract Students to Science Who Want to Solve Real-World Problems

This approach may be especially attractive to those students who would otherwise become disenfranchised from science through traditional approaches to teaching and learning. Emerging research is demonstrating that allowing students to make connections between the science they study and the problems that they, their families, and their communities face can encourage greater interest in science as well as the motivation to learn scientific concepts more deeply (NRC, 2000b; Hulleman and Harackiewicz, 2009).

The New Biologist Is Not a Scientist Who Knows a Little about All Disciplines, but One with Deep Knowledge in One Discipline and a "Working Fluency" in Several

Although this vision of scientists who participate in the New Biology may seem to support the current structure of science majors, it actually would require very different thinking about how scientists are educated. Solving complex, interdisciplinary problems will require that students go far beyond their life science majors both in understanding what connections exist across disciplines and how to make those connections. Requiring separate courses in other natural and behavioral sciences with no attempt to help students make specific connections among them will probably be insufficient. Preparing future life scientists without offering them exposure to and experience with engineering, design, computer science, and an appreciation of the broader connections between science and technology (NRC, 1998, 2003; National Academy of Engineering, 2002, 2007, 2009) will not constitute adequate preparation. And mere exposure (by requiring students to take courses in these other areas) most likely will not prepare them to make and understand the connections among these disciplines; specific efforts must be made to help students learn these skills (NRC, 2000b).

Highly Developed Quantitative Skills Will Be Increasingly Important

Mathematics and other quantitative tools are becoming increasingly important to the work of biologists and to the advancement of the field, and these areas need to become a larger part of undergraduate biology education (NRC, 2003a; Bialek and Botstein, 2004[5]; Brent, 2004; Cohen, 2004; Hoy, 2004; Gross, 2004; Steen, 2005). However, there are many structural and systemic impediments that limit true integration of mathematics and quantitative literacy into undergraduate biology education. These include lack of communication between biology and mathematics departments to better integrate mathematical concepts and examples into biology courses and more appropriate examples involving biology in mathematics courses in which biology majors enroll. There are also persistent misperceptions about the kinds of mathematics that are required to prepare premedical students for the Medical College Admission Test (currently none are specifically required[6]) or for entrance to medical schools (requirements vary widely from urging preparation in mathematics, to one or two semesters of calculus or to algebra or statistics[7]). The recent report from the Association of American Medical Colleges and HHMI (2009) recommends that students should be able to "Apply quantitative reasoning and appropriate mathematics to describe or explain phenomena in the natural world." (p. 22). This competency could be demonstrated by students who are able to

- demonstrate quantitative numeracy and facility with the language of mathematics,
- interpret data sets and communicate those interpretations using visual and other appropriate tools,
- make statistical inferences from data sets,
- extract relevant information from large data sets,
- make inferences about natural phenomena using mathematical models,
- apply algorithmic approaches and principles of logic (including the distinction between cause/effect and association) to problem-solving,
- quantify and interpret changes in dynamical systems (pp. 22–24).

[5] This editorial is part of a special issue of *Science* on "Mathematics in Biology." All relevant papers in this issue are available through links at www.sciencemag.org/sciext/mathbio.

[6] According to the Association of American Medical Colleges, "The Medical College Admission Test (MCAT) is a standardized, multiple-choice examination designed to assess the examinee's problem solving, critical thinking, writing skills, and knowledge of science concepts and principles prerequisite to the study of medicine. Scores are reported in Verbal Reasoning, Physical Sciences, Writing Sample, and Biological Sciences. Medical colleges consider MCAT exam scores as part of their admission process." See www.aamc.org/students/mcat/about/start.htm.

[7] For a listing of entry requirements in mathematics for medical schools in the United States, see www.cse.emory.edu/sciencenet/additional_math_reqs.pdf.

New thinking about ways to integrate and connect these two disciplines can serve as the basis for departments of biology and mathematics, and for professional societies in these disciplines, to work together toward the improvement of undergraduate education as envisioned by the New Biology.

Development and Implementation of Genuinely Interdisciplinary Undergraduate Courses and Curricula Will Both Prepare Students for Careers as New Biology Researchers and Educate a New Generation of Science Teachers Who Will Be Well Versed in New Biology Approaches

The preparation of future science teachers must become a joint responsibility between faculties in science departments and schools of education (NRC, 1998, 2000a, 2003a). Templates and syllabi for interdisciplinary undergraduate courses that would benefit teachers of science (especially those in the elementary and middle grades) have been published.[8] But science, mathematics, and engineering faculty and academic leaders in higher education must recognize their roles in preparing future teachers as well as future researchers. Consideration must be given to what undergraduates will need to learn to teach science in the way envisioned in *A New Biology*, both with respect to the necessary scientific knowledge base and to familiarity with scientifically based pedagogical techniques that are most effective in teaching science.

Similar attention needs to be paid to preparing graduate students to become the next generation of faculty who will, in turn, assume some of the responsibility for K–12 teacher preparation. Are graduate students being encouraged to pursue quality teaching experiences? Are they being provided with training in new approaches to teaching and learning and exposure to the research literature about human learning and cognition as part of that preparation?

What characteristics might undergraduate courses have that emphasize an interdisciplinary approach as envisioned in *A New Biology*? The report provides an example of introductory courses at Harvard University (see Box 2). Additional models are offered by SENCER (see footnote 9) and include courses with biological emphases such as

- Cellular and Molecular Biology: Cancer
- Life Science in Context: SubSaharan Africa & HIV/AIDS
- The Science of Sleep
- Slow Food
- Addiction: Biology, Psychology, and Society
- Environment and Disease
- Nutrition & Wellness and the Iowa Environment
- Human Genetics
- Tuberculosis
- Biomedical Issues of HIV/AIDS
- Mysteries of Migration

[8] For example, model courses have been developed with support from the NSF as part of the Science Education for New Civic Engagements and Responsibilities (SENCER); see www.sencer.net/Resources/models.cfm) and the Mathematics/Science Partnerships (see http://mspnet.org) initiatives.

Box 1: Connecting *Bio 101* to Real-World Issues: An Interdisciplinary Approach

In 2005–2006, Harvard University launched two semester-long introductory courses that provide an interdisciplinary introduction to biology and chemistry. The first course synthesizes essential topics in chemistry, molecular biology, and cell biology, and the second course synthesizes essential topics in genetics, genomics, probability, and evolutionary biology. Scientific facts and concepts are introduced in the context of exciting and interdisciplinary questions, such as understanding the possibility of synthetic life, the biology and treatment of AIDS and cancer, human population genetics, and malaria. Through interdisciplinary teaching, students' grasp of fundamental concepts is reinforced as they encounter the same principles in multiple situations. Each course is taught by a small team of faculty from multiple departments. Members of each teaching team attend all lectures and participate for the entire term. The preparation for and teaching effort in each course offering is integrated. Teaching assistants are also drawn from different departments and work in small interdepartmental teams.

Development of these courses required institutional support. The president, dean of the faculty, and the chair of the life sciences council all provided funds to support a one-year curriculum development effort, lab renovations, lower teaching fellow–student ratios, equipment, and development of teaching materials. One of the founding faculty member's HHMI undergraduate education award contributed to developing specific sets of teaching materials.

Success depended on finding faculty members with personal commitments to the principles of the courses and willingness to work as a team to build the new courses from scratch. This effort was rewarded as individual departments agreed to count these interdepartmental and interdisciplinary courses toward their respective departmental teaching expectations.

Since the courses were implemented, undergraduate enrollment in introductory life sciences courses is up >30% and the number of life sciences majors has risen 18%. NRC (2009) p. 80

The life sciences and science education communities have made significant advances in articulating how undergraduate biology education can be made accessible to more students with varying education needs and learning styles. The beginnings of real consensus about the future course for life sciences education is emerging. As the year 2010 opens, the ideas for "transforming undergraduate education for future research biologists" that were envisioned in the *Bio 2010* report are being considered more seriously and implemented more widely than many had imagined when the report was published in 2003 (e.g., Pfund *et al*., 2009). The *New Biology* report emphasized the ongoing and lasting relevance of *Bio 2010* but also noted the incomplete implementation of its recommendations to date. Much work remains.

The findings and recommendations that emerged in 2009 again offer a collective and coherent vision for improving undergraduate science education in general, and biology education specifically. As a community, we must work toward implementation of the visions articulated in *A New Biology* and other recent initiatives, scaled to encompass all areas of biology and all undergraduates who enroll in biology courses and programs.

REFERENCES

Allen, D., and Tanner, K. (2007). Putting the Horse Back in Front of the Cart: Using Visions and Decisions about High-Quality Learning Experiences to Drive Course Design. CBE Life Sci. Educ. *6*, 85–89. www.lifescied.org/cgi/content/full/6/2/85?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=&fulltext=human+learning&andorexactfulltext=and&searchid=1&FIRSTINDEX=0&sortspec=relevance&resourcetype=HWCIT (accessed 9 December 2009).

Association of American Medical Colleges and Howard Hughes Medical Institute (2009). Scientific Foundations for Future Physicians: Report of the AAMC-HHMI Committee, Washington, DC, and Chevy Chase, MD. www.hhmi.org/grants/pdf/08-209_AAMC-HHMI_report.pdf (accessed 5 December 2009).

Bialek, W., and Botstein, D. (2004). Introductory science and mathematics education for 21st-century biologists. Science *303*, 788–790. www.sciencemag.org/cgi/content/full/303/5659/788 (accessed 9 December 2009).

Brent, R. (2004). Points of view: the interface of mathematics and biology: intuition and innumeracy. Cell Biol. Educ. *3*, 88–90. www.lifescied.org/cgi/content/full/3/2/88 (accessed 9 December 2009).

Cohen, J. E. (2004). Mathematics is biology's next microscope, only better; biology is mathematics' next physics, only better. PLoS Biol. *2*, e439. www.plosbiology.org/article/info%3Adoi%2F10.1371%2Fjournal.pbio.0020439 (accessed 9 December 2009).

DeHaan, R. L. (2009). Teaching creativity and inventive problem solving in science. CBE Life Sci. Educ. *8*, 172–181. www.lifescied.org/cgi/content/full/8/3/172?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=&fulltext=human+cognition&andorexactfulltext=and&searchid=1&FIRSTINDEX=0&sortspec=relevance&resourcetype=HWCIT (accessed 9 December 2009).

Gross, L. J. (2004). Points of view: the interface of mathematics and biology: interdisciplinarity and the undergraduate biology curriculum: finding a balance. Cell Biol. Educ. *3*, 85–87. www.lifescied.org/cgi/content/full/3/2/85 (accessed 9 December 2009).

Handelsman *et al*. (2006). Scientific Teaching, San Francisco, CA: Freeman and Co.

Hoy, R. (2004). Points of view: the interface of mathematics and biology: new math for biology is the old math. Cell Biol. Educ. *3*, 90–92. www.lifescied.org/cgi/content/full/3/2/90 (accessed 9 December 2009).

Hulleman, C., and Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. Science *326*, 1410–1412.

Jurkowski, A., Reid, A. H., and Labov, J. B. (2007). Metagenomics: a call for bringing a new science into the classroom (while it's still new). CBE Life Sci. Educ. *6*, 260–265. www.lifescied.org/cgi/content/full/6/4/260?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=&author1=jurkowski&andorexactfulltext=and&searchid=1&FIRSTINDEX=0&sortspec=relevance&resourcetype=HWCIT (accessed 5 December 2009).

Labov, J. B., Singer, S. R., George, M. D., Schweingruber, H. A., and Hilton, M. L. (2009). From the national academies: effective practices in undergraduate STEM education part 1, examining the evidence. CBE Life Sci. Educ. *8*: 157–161. www.lifescied.org/cgi/content/full/8/3/157?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=1&author1=labov&andorexacttitle=and&andorexacttitleabs=and&andorexactfulltext=and&searchid=1&FIRSTINDEX=0&sortspec=relevance&resourcetype=HWCIT,HWELTR (accessed 13 December 2009).

Mervis, J. (2009a). Revisions to AP courses expected to have domino effect. Science *325*, 1488–1498. www.sciencemag.org/cgi/reprint/325/5947/1488.pdf (accessed 10 December 2009).

Mervis, J. (2009b). Universities begin to rethink university biology courses. Science *325*, 527. www.sciencemag.org/cgi/reprint/325/5940/527-a.pdf (accessed 10 December 2009).

Morse, D., and Jutras, F. (2008). Implementing concept-based learning in a large undergraduate classroom. CBE Life Sci Educ. *7*, 243–253. www.lifescied.org/cgi/content/full/7/2/243?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=&fulltext=human+cognition&andorexactfulltext=and&searchid=1&FIRSTINDEX=0&sortspec=relevance&resourcetype=HWCIT (accessed 9 December 2009).

National Academy of Engineering (NAE) (2002). Technically Speaking: Why All Americans Need to Know More About Technology, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=10250 (accessed 9 December 2009).

NAE (2007). Exploring Opportunities in Green Chemistry and Engineering Education: A Workshop Summary to the Chemical Sciences Roundtable, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=11843 (accessed 9 December 2009).

NAE (2009). Engineering in K–12 Education: Understanding the Status and Improving the Prospects, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=12635 (accessed 9 December 2009).

National Research Council (NRC) (1996). From Analysis to Action: Undergraduate Education in Science, Mathematics, Engineering, and Technology, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=9128 (accessed 5 December 2009).

NRC (1997). Science Teaching Reconsidered: A Handbook, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=5287 (accessed 5 December 2009).

NRC (1998). Transforming Undergraduate Education in Science, Mathematics, Engineering, and Technology, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=6453 (accessed 5 December 2009).

NRC (2000a). Educating Teachers of Science, Mathematics, and Technology: New Practices for the New Millennium, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=9832 (accessed 5 December 2009).

NRC (2000b). How People Learn: Brain, Mind, Experience, and School: Expanded Edition, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=9853 (accessed 5 December 2009).

NRC (2002). Learning and Understanding: Improving Advanced Study of Mathematics and Science for U.S. High School Students, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=10129 (accessed 5 December 2009).

NRC (2003a). Bio 2010, Transforming Undergraduate Biology Education for Future Research Biologists, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=10497 (accessed 5 December 2009).

NRC (2003b). Evaluating and Improving Undergraduate Education in Science, Technology, Engineering, and Mathematics, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=10024 (accessed 5 December 2009).

NRC (2003c). Improving Undergraduate Instruction in Science, Technology, Engineering, and Mathematics: Report of a Workshop, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=10711 (accessed December 5, 2009).

NRC (2005). America's Lab Report: Investigations in High School Science, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record[lwoen]id=11311 (accessed 5 December 2009).

NRC (2008). The Role of Theory in Advancing 21st-Century Biology: Catalyzing Transformative Research, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=12026 (accessed 5 December 2009).

NRC (2009). A New Biology for the 21st Century: Ensuring the United States Leads the Coming Biology Revolution, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=12764 (accessed 5 December 2009).

National Science Foundation (1996). Shaping the Future: New Expectations for Undergraduate Education in Science, Mathematics, Engineering, and Technology, Arlington, VA: National Science Foundation. www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf96139 (accessed 5 December 2009).

Pfund, C., *et al*. (2009). Summer institute to improve biology education at research universities. Science *324*, 470–471. www.sciencemag.org/cgi/content/summary/324/5926/470 (accessed 9 December 2009).

Steen, L. A. (ed.) (2005). Math and Bio 2010, Linking Undergraduate Disciplines, Washington, DC: Mathematical Association of America. http://books.google.com/books?hl=en&lr=&id=JjK0bqx6Xa8C&oi=fnd&pg=PR7&dq=%22Steen%22+%22Math+and+bio+ *2010*, +linking+undergraduate+disciplines%22+&ots=mGNta2SBi_&sig=M7WCVFHmN0osVmXY1sva0bI4YPQ#v=onepage&q=&f=false (accessed 9 December 2009).

Wood, W. B. (2009). Revising the AP biology curriculum. Science *325*, 1627–1628. www.sciencemag.org/cgi/reprint/325/5948/1627.pdf (accessed 10 December 2009).

Woodin, T., Smith, D., and Allen, D. (2009). Transforming undergraduate biology education for all students: an action plan for the twenty-first century. CBE Life Sci. Educ. *8*, 271–273. www.lifescied.org/cgi/content/full/8/4/271#B12 (accessed 5 December 2009).

CBE—Life Sciences Education
Vol. 7, 347–352, Winter 2008

Feature
From the National Academies

Integrating Policy and Decision Making into Undergraduate Science Education

Jay B. Labov* and Nancy F. Huddleston†

*Center for Education and †Division of Earth and Life Sciences, National Research Council, Washington, DC 20001

INTRODUCTION

> *Scientists gathered in Mill Valley [CA] Thursday as part of a fact-finding mission to determine what effect the Drakes Bay Oyster Co. has on the ecology of Drakes Estero. The company's lease allowing it to grow and harvest oysters in Drakes Estero ends in 2012, and the Point Reyes National Seashore wants to turn it into a wilderness area thereafter. But owner Kevin Lunny said the operation causes no harm and may help the ecosystem. He wants to stay.*
>
> *The National Research Council—an arm of the National Academy of Sciences—was tapped by the National Park Service to examine the issue at the request of Sen. Dianne Feinstein, D.-Calif. That process began Thursday as the nine members of the committee—including experts in agriculture, disease, marine sciences and oceanography—heard from a variety of people connected to the issue in what had the feel of a courtroom at the Aqua Hotel . . .*
>
> Excerpted from Prado, 2008.

> *"There is no end of examples of policies that have been established at the state level that have failed dramatically because they have not taken into account science and technology issues."*
>
> Richard Atkinson, President Emeritus, University of California. In National Academy of Science, National Academy of Engineering, and Institute of Medicine [NAS, NAE, IOM], 2008, p. 2.

The above-mentioned quotes offer important messages and frameworks for readers of *CBE—Life Sciences Education*. First, the need for and impact of science and solid scientific evidence can be found in virtually all aspects of life. Second, science can and should inform policy and decision making to a much greater extent than is currently the case. Third, the science that we teach and ask our students to learn can be infused with real-life examples of policy issues that can help many more students, whether they plan to go on to careers in science or not, to understand science in context. Fourth, although Atkinson's statement notes that many policies have failed because they have not been infused with scientific and technological information and perspective, the majority of people who are elected or appointed as policy makers have taken college-level science courses. However, too many of them have not been prepared adequately to deal with the implications of science in policy and decision making. Even those policy makers who may be steeped in the content of a particular science discipline (i.e., science majors) may not have been asked to understand or explore deeply the processes, nature, and limits of science in decision making.

It is clear that most students who graduate from college will not enter the public policy arena as part of their careers. Still, too many of them leave school without making the connections about how scientific and technological concepts relate to a myriad of issues that affect their daily lives. These include their health and the well-being of themselves and their families, and their roles and responsibilities as citizens of their local community, their nation, and an increasingly interconnected, interdependent, and scientifically driven world. Because the majority of undergraduates do not enroll in science courses beyond the introductory level, infusing these kinds of ideas and perspectives into those courses is especially important (e.g., National Research Council [NRC], 1999; Labov, 2004; Jurkowski *et al*., 2007). Doing so can improve the course at several levels, increasing student participation, enthusiasm, and knowledge retention, as discussed in the fall issue of *CBE−Life Sciences Education* (Chamany *et al*., 2008).

A workshop that the National Academies convened in 2007 focused on the problems in using science and technology information to guide public policy and decision making at the state level. Today, state leaders are responsible for an increasing number of decisions that rely on access to high-quality scientific information—from creating their state's technology portfolio, to managing water and other resources, maintaining critical infrastructure, and improving education and health care. The workshop brought together state leaders with the many sources of

DOI: 10.1187/cbe.08–09–0056
Address correspondence to: Jay B. Labov (jlabov@nas.edu).

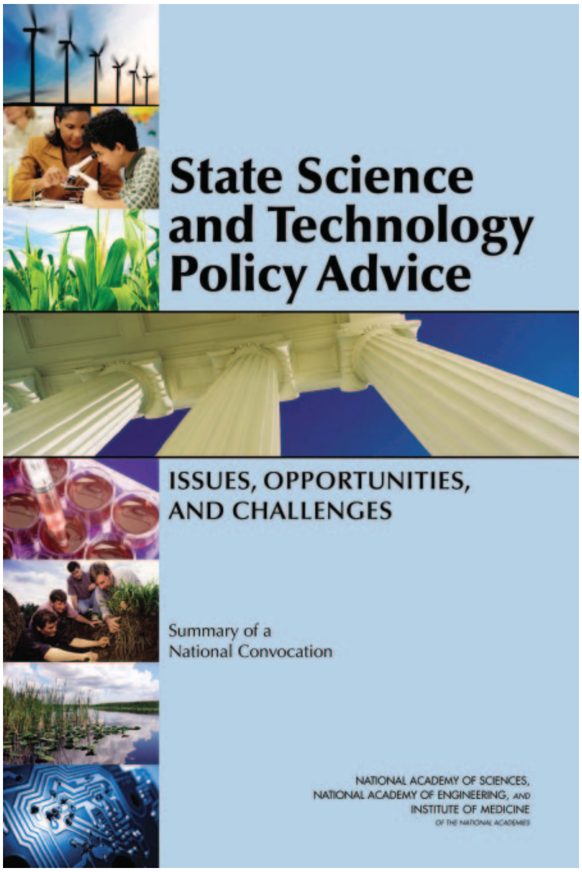


Figure 1. Cover of NAS, NAE, IOM (2008) summary of the national convocation on state science and technology policy advice.

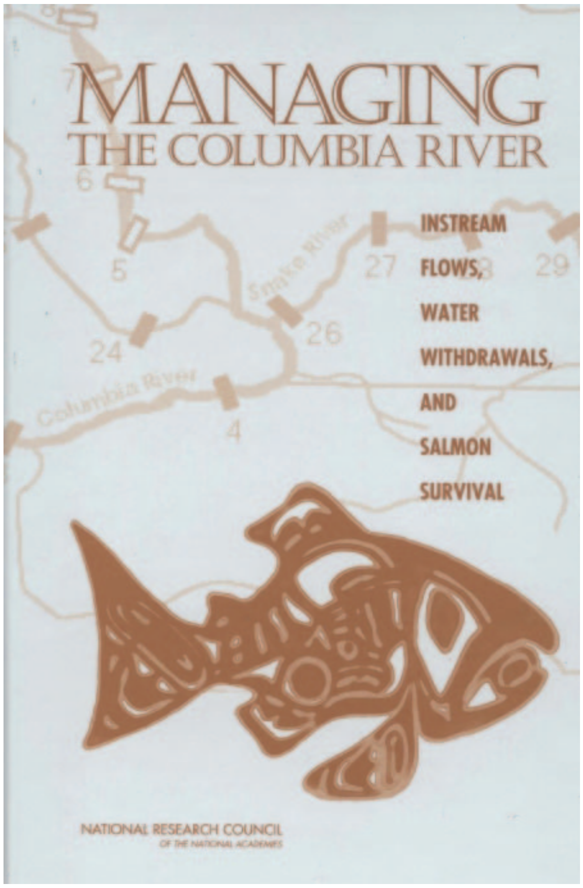


Figure 2. Cover of NRC (2004a) report on Managing the Columbia River.

science advice (NAS, NAE, IOM, 2008[1]; the cover of this report is displayed in Figure 1.) The workshop pointed both to the need for decision makers to pay more attention to science guidance and for scientists to improve the content and timing of their communications.

THE WASHINGTON STATE STORY

At the October 2007 convocation (NAS, NAE, IOM, 2008), Gerry O'Keefe, Columbia River policy coordinator for the state of Washington's Department of Ecology, provided an excellent example of one state's struggle to get and use science advice to break through a decision-making impasse in how to manage the precious water resources of the Columbia River Basin (NRC, 2004a; Figure 2). The details of this case that are presented below are excerpted and modified from the report from that convocation.

The Columbia River carries 200 million acre-feet of water in an average year (which, coincidentally, is about the same size as the water budget for the state of California, O'Keefe noted). It drains an area of 273,000 square miles that extends from Canada to Wyoming and Utah. It is a tightly controlled system that is managed for flood control, agriculture, power generation, and for protection of the salmon that live and spawn in the river, which have iconic value to the people of Washington state. Factors affecting the river are undergoing profound changes, O'Keefe pointed out. Population growth is increasing the demands being made on the river. Climate change, particularly as it affects mountain snowpacks, could alter the amount of water that the river can supply. Salmon species in the river are in decline.

However, the river continues to offer untapped potential for economic development. According to one calculation, withdrawing just 1 million acre-feet of water, which is about half of 1% of the river's annual flow, and applying it to the land would create 18,000 jobs and annual revenues of approximately $850 million. "This is a number that is not ever ignored by the governor's office—or the state legislature," said O'Keefe. "It captures and crystallizes their attention like almost nothing else will."

[1] These three organizations plus the NRC comprise the National Academies, a private, nonprofit organization whose charter is to assist the federal government and the American people in analyzing pressing science and technology policy issues. The NRC serves as the operating and research arm. Much of the work of the National Academies is performed by leading scientists, mathematicians, engineers, social scientists, and policy experts who provide pro bono service to the National Academies and the nation. More information about the National Academies is available at http://nationalacademies.org.

For decades, the state has struggled to develop policies to manage the Columbia River Basin. Many groups have conflicting interests in the Columbia River, including farmers, manufacturers, other private interests, the federal government, the environmental community, and 13 Native American tribes that rely on the river's water. As discussions among these groups deteriorated over the years, management decisions became increasingly difficult. "When state officials or others in charge of mediating among the sides tried to arrange meetings, the sides would not even agree to talk unless they knew what the outcome of the discussion was likely to be." Different groups "have veto power," said O'Keefe. "The federal statute is designed with overlapping authorities and jurisdictions, and unless you have something close to consensus, you're going to find out that you're unable to act."

In 2002, the state turned to the Water Science and Technology Board at the NRC[2] for help. The first task was to define the question to be addressed. "We spent a tremendous amount of time and energy thinking about what it was we were going to ask the National Academy of Sciences [NRC] to resolve for us." The actual charge covered most of two pages, but it can be boiled down to a relatively simple question: If 1 million acre-feet of water were to be removed from the river, what impact would that action have on endangered species, and what could be done to mitigate those impacts? The state did not know what the response from the NRC committee would be, and the final report (NRC, 2004a) did not deliver the answer that the state expected, according to O'Keefe. Although state officials expected that a relatively small withdrawal of water from the river was unlikely to have a measurable effect on the salmon, the NRC report said otherwise. Instead, the committee concluded that salmon populations were in trouble, especially during summers when the flow of the river is lower and the water is warmer. The conclusion of the report, said O'Keefe, was that "you need to be very careful as you allocate water out of the stream. You are getting yourself into a situation where you could end up with a year or a series of years where you have lost your management flexibility and you have in fact predetermined that you will lose your species as well."

Once the report was delivered, policy makers in Washington state had to decide what to do with the NRC's advice. This was not a foregone conclusion, said O'Keefe. State legislators "really are representative of the communities that elect them. They come from all kinds of backgrounds . . . Our challenge is to try to find ways to . . . connect with those people who have the ability to make those decisions." To their credit, despite the many other competing pressures exerted on them, the state's policy makers did not ignore the advice. "We tried, to the extent we could, to be guided by the National Academies to create a flexible and responsive policy framework on the fly that helped us break through the policy gridlock that we had experienced as a state."

State officials opted to look at additional storage developments for Columbia River water and at the use of existing storage facilities. Of every three quantities of water made newly available through this process, one would be set aside for protection of the salmon. O'Keefe reported that the state linked economic and long-term environmental interests of the state in ways that are very creative, and the result turned out to be quite compelling and powerful. Legislation passed in 2006 authorized the creation of a new water program supported by $200 million of funding to develop water supplies over time. And conversations with officials from Canada and surrounding states were initiated to manage the river more effectively. "The future in Washington state as a result of this conversation is really quite a lot brighter," O'Keefe concluded.

IMPORTANCE OF AND RESOURCES FOR UNDERSTANDING SCIENCE AND POLICY

Introducing public policy concepts and perspectives into science courses may seem daunting to some faculty. Few have been explicitly educated to teach science from such a perspective. However, this is a critical component of a well-rounded science education. Many students, especially nonscience majors, are likely to become more interested in science when they can see the relevance of the subject to other things about which they are interested or even impassioned. And there are increasing numbers of resources to help faculty become better versed in the intersection between science, technology, and public policy and decision making.

For example, over the past 6 yr, the National Science Foundation has supported (through its SENCER[3] initiative) the development of 37 model undergraduate courses that explicitly connect science, technology, engineering, and mathematics with "capacious" questions related to public policy and civic engagement. Many of these courses have a local- or state-level focus. A list of these model courses and links to their descriptions is available at www.sencer.net/Resources/models.cfm.

Similarly, the recently published *Pathways to Scientific Teaching* (Ebert-May and Hodder, 2008) takes articles about ecology published over the years through the Ecological Society of America and helps faculty "translate" them into various kinds of exercises for use in undergraduate classrooms and laboratories. Ebert-May and Hodder (2008) builds on *Scientific Teaching* (Handelsman *et al*., 2004, 2007), which offers faculty concrete and evidence-based methods for incorporating research principles into undergraduate courses.

As described above, the National Academies also can serve as a rich source of resources for engaging students in the intersections between science, technology, and public policy. As noted in the first quote, study committees from the Academies' NRC are asked to examine some of the most difficult policy questions facing society today and to inform Congress, federal and state departments and agencies, and others how scientific and technological evidence can contribute to the development and implementation of policy based on what is currently known.

[2] Additional information about the Water Science and Technology Board is available at http://dels.nas.edu/wstb/.

[3] SENCER: Science Education through New Civic Engagements and Responsibilities. Additional information about SENCER is available at www.sencer.net.

Equally important, Academy committees also help inform their sponsors and broader audiences about what is *not* currently known or well understood. Students may be inspired by the challenge of addressing ongoing issues by developing the scientific context.

Below, we provide several brief synopses of reports that various committees of the NRC have authored for use by states that also could provide compelling examples to students about how science contributes to knowledge and decision making. The National Academies also have worked hard in recent years to develop a variety of products that are derived from our reports that could serve as useful supplements to undergraduate courses in biology and related disciplines; we also discuss how to access these resources.

Assessing the Safety of a Biocontainment Lab in Boston

At the request of the state of Massachusetts, the NRC reviewed a draft document from the National Institutes of Health (NIH) assessing the risks of a new National Emerging Infectious Diseases Laboratory being built at Boston University. The facility would include a Biosafety Level 4 laboratory for research on deadly pathogens such as the Ebola virus. The opening of the lab had been challenged through both state and federal lawsuits.

The committee report, *Technical Input on the National Institutes of Health's Draft Supplementary Risk Assessments for the National Emerging Infectious Diseases Laboratory, Boston University* (NRC, 2007a) concluded that the NIH draft risk assessment report has serious weaknesses and does not adequately identify, or thoroughly develop, worst case scenarios for the release and spread of a pathogen. The report commends the NIH for working with the community to identify pathogens to include in the scenarios, but finds that the process seems to have led to the selection of pathogens that do not fully address matters raised by the state. The report concludes that NIH should have included agents that are readily transmissible and would have demonstrated that the modeling approach used recognizes biological complexities, reflecting what is known about disease outbreaks and being appropriately sensitive to population density, for example. This report raises many issues of interest for anyone teaching a microbiology course, particularly a lab course, as well as addressing interesting issues in public health.

Oysters in the Chesapeake Bay

Long before the Drakes Bay Oyster dispute that was mentioned at the beginning of this article, the states of Maryland and Virginia faced their own oyster problem. Decades of heavy fishing, environmental pressures, and deadly disease have nearly eradicated native oysters in the Chesapeake Bay and a once-thriving oyster industry. Because oysters feed on algae, their disappearance is thought to play a role in the general decline of water quality in the Bay, which often becomes algae-laden during parts of the year. At the request of the Maryland and Virginia and federal partner agencies, the National Academies identified potential risks and benefits of introducing the Asian suminoe oyster to supplement or replace the disease-plagued native species. Opponents feared that the nonnative species could become invasive, with potentially devastating impacts on the ecology and economic vitality of the region.

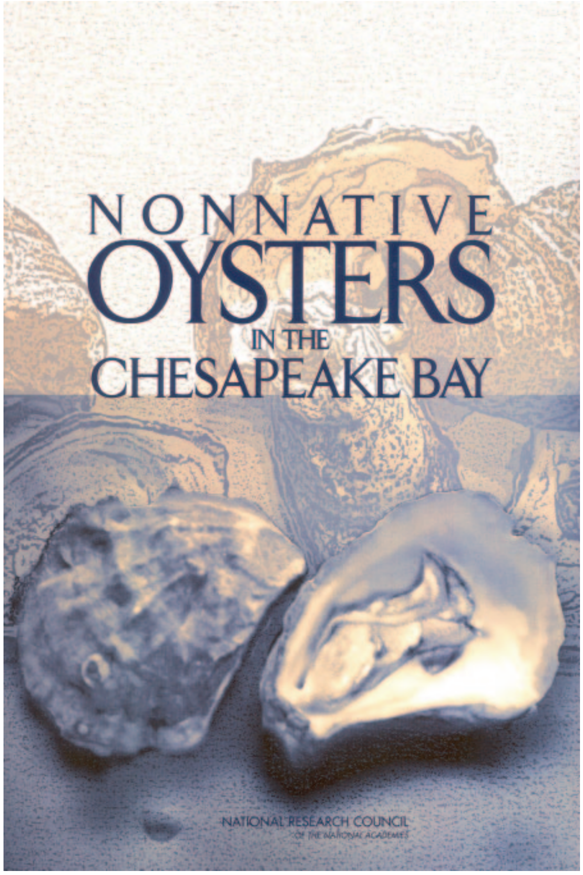


Figure 3. Cover of NRC (2004b) report on the Chesapeake Bay oyster proposal.

Nonnative Oysters in the Chesapeake Bay (NRC, 2004b; Figure 3) recommends aquaculture of nonreproductive suminoe oysters as the most prudent option until completion of research to investigate the potential impacts of introducing this nonnative oyster species. The report also proposes stricter regulations to reduce the risk of unintentional introductions of nonnative species. These issues are relevant for courses in ecology at any level.

Guiding Stem Cell Research in California

In 2004, the State of California sought advice from the National Academies about how to create major new programs that voters had approved for state-funded stem cell research. To help guide the state in its research planning, the National Academies convened experts in the field for a 2-day workshop in California. Topics discussed at the workshop included grant-making processes, intellectual property, institutional review boards, facility development, and the development of standards and ethical

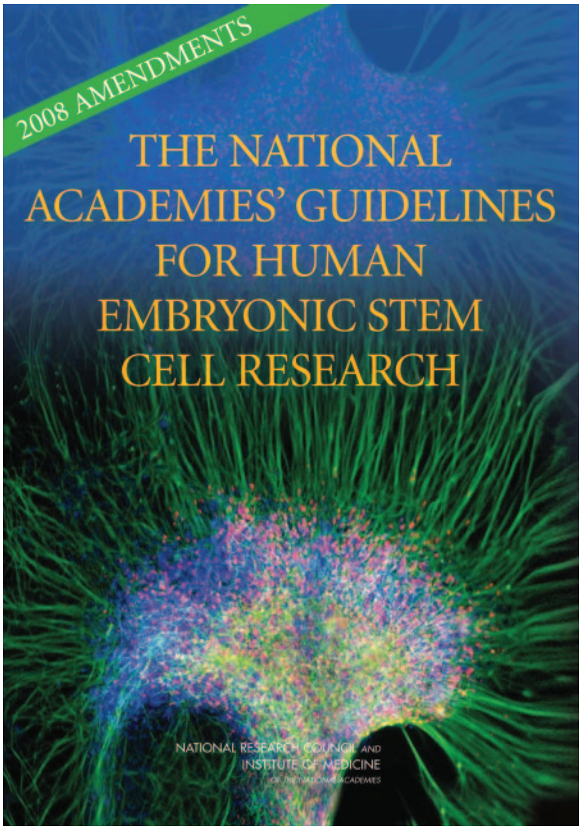


Figure 4. Cover of NRC (2008) guidelines for human embryonic stem cell research.

guidelines. The ethical standards discussed at the workshop were the result of other work the Academies were doing at the time to develop its *Guidelines for Human Embryonic Stem Cell Research* (NRC, 2005). This report recommended the establishment of an oversight system for human embryonic stem cell research that has been widely adopted nationwide. The original report has since been updated twice (NRC 2007b, 2008; see also Figure 4). Asking students to view the progression of these guidelines can offer great insight into the rapid progress of cutting-edge science and its implications for public policy and welfare. The basic science underlying this topic derives from studies in developmental biology, cell biology, and gene regulation, and can be approached from a relatively simple or a very sophisticated level of instruction.

Watershed Solutions in New York

The state of New York had always enjoyed high-quality water from the Catskill Mountain watershed, which provides ~90% of the drinking water for New York City. Unfortunately, increased numbers of housing developments and associated septic systems, and the impacts of agriculture, have caused water quality to deteriorate. By the late 1990s, New York City water managers had two choices: build a water filtration system at an estimated cost of up to $6 billion or take steps to protect its major watershed.

To help weigh the scientific and technical aspects of its dilemma, the state turned to the National Academies. On the basis of recommendations in *Watershed Management for Potable Water Supply: Addressing the New York City Strategy* (NRC, 2000), stakeholders decided against building the filtration system and began taking recommended steps to protect the watershed at a total projected investment of ~$1–1.5 billion. Water use and quality is, of course, a central topic for any course on natural resources or sustainability and can be introduced in courses on microbiology, public health, and ecology.

ACCESSING NATIONAL ACADEMIES RESOURCES

Readers may access information about National Academies studies in various ways. Information about all studies in progress can be searched through the Academies' Current Projects System.[4]

Additional information about current projects and completed projects can be obtained by entering keywords into the search engine or by selecting a general disciplinary area located on the left side of the National Academies' home page.[5] Each of the disciplinary divisions within the Academy have worked hard over the past 10 years to make findings from their reports accessible to broader audiences through a series of derivative products. For example, a listing of an extensive series of such resources in the earth and life sciences can be found at http://dels.nas.edu/dels/sp_products.shtml.

The National Academies Press provides several ways to search for Academies' reports and related information.[6] A very sensitive search engine allows users to enter parts of the title of a report or key words to locate all reports that the Academies have published on this topic. This website also contains several other features that readers can use to find information:

- After a particular report has been located, readers can use the Web Search Builder tool to use key words or phrases from that book to search for information within that book, across the academies' collection of resources, or across the Web.
- The Reference Finder allows readers to paste in their own text to find books that are related to the topic. For example, a sentence containing key words or phrases from a journal article or even from a student's draft term paper can be entered to search for additional information.

We hope that these examples are illustrative. Resources are available as described above and through many other sources (see Chamany *et al.*, 2008). Opportunities to make science and technology more relevant to many more students await!

[4] Available at www8.nationalacademies.org/cp.
[5] Available at http://nationalacademies.org.
[6] Available at http://nap.edu.

REFERENCES

Chamany, K., Allen, D., and Tanner, K. (2008). Making biology learning relevant to students: integrating people, history, and context into college biology teaching. CBE Life Sci. Educ. *7*, 267–278. http://www.lifescied.org/cgi/reprint/7/3/267 (accessed 14 October 2008).

Ebert-May, D., and Hodder, J. (eds.) (2008). Pathways to Scientific Teaching, Sunderland, MA: Sinauer Associates, Inc.

Handelsman, J. *et al*. (2004). Scientific teaching. Science *30*, 521–522. http://www.sciencemag.org/cgi/content/full/304/5670/521?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=&fulltext=handelsman&searchid=1109276347970_1405&stored_search=&FIRSTINDEX=0 (accessed 14 October 2008).

Handelsman, J., Miller, S., and Pfund, C. (2007). Scientific Teaching, New York: W.H. Freeman and Company.

Jurkowski, A., Reid, A. H., and Labov, J. B. (2007). Metagenomics: a call for bringing a new science into the classroom (while it's still new). CBE Life Sci. Educ. *6*, 260–265. www.lifescied.org/cgi/content/full/6/4/260?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=1&author1=labov&andorexacttitle=and&andorexacttitleabs=and&andorexactfulltext=and&searchid=1&FIRSTINDEX=0&sortspec=relevance&resourcetype=HWCIT,HWELTR (accessed 14 October 2008).

Labov, J. B. (2004). From the National Academies: the challenges and opportunities for improving undergraduate science education through introductory courses. Cell Biol. Educ. *3*, 212–214. www.lifescied.org/cgi/content/full/3/4/212?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=1&author1=labov&andorexacttitle=and&andorexacttitleabs=and&andorexactfulltext=and&searchid=1&FIRSTINDEX=10&sortspec=relevance&resourcetype=HWCIT,HWELTR (accessed 14 October 2008).

National Academy of Sciences, National Academy of Engineering, and Institute of Medicine (NAS, NAE, IOM). 2008. State Science and Technology Policy Advice: Issues, Opportunities, and Challenges. Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=12160 (accessed 14 October 2008).

National Research Council (NRC). 1999. Transforming Undergraduate Education in Science, Mathematics, Engineering, and Technology, Washington, DC: National Academies Press. http://nap.edu/catalog/6453.html (accessed 14 October 2008).

NRC. 2000. Watershed Management for Potable Water Supply: Addressing the New York City Strategy, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=9677 (accessed 14 October 2008).

NRC. 2004a. Managing the Columbia River: Instream Flows, Water Withdrawals, and Salmon Survival, Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=10962 (accessed 14 October 2008).

NRC. 2004b. Nonnative oysters in the Chesapeake Bay, Washington DC: National Academies Press. www.nap.edu/catalog.php?record_id=10796 (accessed 14 October 2008).

NRC. 2005. Guidelines for Human Embryonic Stem Cell Research, Washington DC: National Academies Press. www.nap.edu/catalog.php?record_id=11278#description (accessed 14 October 2008).

NRC. 2007a. Technical Input on the National Institutes of Health's Draft Supplementary Risk Assessments and Site Suitability Analyses for the National Emerging Infections Diseases Laboratory, Boston, University: A Letter Report. Washington, DC: National Academies Press. www.nap.edu/catalog.php?record_id=12073 (accessed 14 October 2008).

NRC. 2007b. 2007 Amendments to the National Academies' Guidelines for Embryonic Stem Cell Research, Washington DC: National Academies Press. www.nap.edu/catalog.php?record_id=11871 (accessed 14 October 2008).

NRC. 2008. 2008 Amendments to the National Academies' Guidelines for Embryonic Stem Cell Research, Washington DC: National Academies Press. www.nap.edu/catalog.php?record_id=12260 (accessed 14 October 2008).

Prado, M. 2008. Impact of oyster harvest at Drakes Estero studied. *Marin Independent Journal*, September 4. www.marinij.com/marinnews/ci_10382590?IADID=Search-www.marinij.com-www.marinij.com (accessed 14 October 2008).

Sharon Long is Steere-Pfizer Professor of Biological Sciences and dean emerita of Humanities and Sciences at Stanford University. E-mail: srl@stanford.edu

Robert Alpern is dean and Ensign Professor at Yale University School of Medicine. E-mail: robert.alpern@yale.edu

Science for Future Physicians

BIOMEDICAL RESEARCH IS RAPIDLY TRANSFORMING OUR UNDERSTANDING OF HEALTH AND DISEASE, with major implications for medical practice. But the science education of physicians has not kept pace with these advances. Today, the Association of American Medical Colleges (AAMC) and the Howard Hughes Medical Institute release a report that addresses this issue.* The analysis, by a committee of U.S. undergraduate and medical school faculty that we co-chaired, comes 6 years after the U.S. National Academies report *BIO 2010*, which noted that undergraduate premedical course requirements and the content of the Medical College Admissions Test (MCAT) constrain innovation in undergraduate science education.

The new report, *Scientific Foundations for Future Physicians*, emphasizes that physicians must have a firm grounding in the biomedical sciences and understand their relation to the physical sciences and mathematics. For physicians to be prepared for inquisitive, critical thinking and lifelong learning, they should also be able to incorporate the methods of science into their practice, including skeptical and critical analysis. These goals should be reflected across the entire span of a physician's education, from undergraduate study through medical school.

Medical school faculty have a short time in which to convey an in-depth understanding of specific medical knowledge and recent research. Students should arrive at medical school prepared in the sciences, including some areas not currently required, such as statistics and biochemistry. If all beginning medical students understand general biochemistry, for example, then faculty can build on this knowledge, creating more opportunities to explore the synergistic relationships among biomedical science, research, and clinical medicine. Medical schools should also increase their emphasis on the importance of the physical sciences and mathematics in biomedical research and clinical practice.

How should preparation for medical study be assessed? Medical schools generally determine scientific readiness for admission by course requirements and scores on the MCAT, which mainly reflects the traditional content of those courses. In contrast, medical schools have long evaluated readiness for medical practice in terms of competency—specific learned abilities that can be put into practice—rather than by mandating standard courses and curricula for all medical schools. The report recommends that scientific readiness for medical school entry be assessed similarly: The current list of required premedical school courses should be replaced with required science competencies. Instead of a nationwide requirement that premedical undergraduates take specific chemistry classes, for example, a required competency might be described as being "able to apply knowledge of the chemistry of carbon compounds to biochemical reactions." The report suggests competencies for premedical and medical school science education, recognizing that there may be multiple routes to gaining a competency. An integrated approach to both undergraduate and medical education may help both to innovate.

A change in competency-based educational goals would allow undergraduate institutions the option to design new curricula. Chemistry competencies, for example, might be gained either in traditional chemistry courses or in rigorous interdisciplinary courses. Such innovations, aimed at increasing relevant scientific content and understanding, should also yield more efficient teaching. While recommending new competencies such as biochemistry and statistics, the report's committee opposes a net increase in premedical science requirements. As the report states, "the undergraduate years should not be designed primarily to prepare students for professional school, but for creative engagement in a broad, intellectually expansive education." In conjunction with the shift to competencies, the MCAT could be modified to test for competencies proposed by the committee. The new report coincides with an AAMC review of MCAT content, a multiphase process that will consider the report, among other inputs.

In a system with so many participants, dialogue is essential to progress. Outreach to and feedback from the scientific disciplines and medical community will enhance the success of these efforts. We urge colleagues to engage in this national discussion. **– Sharon Long and Robert Alpern**

*See www.aamc.org/scientificfoundations.

10.1126/science.1176994

Facts and Myths about Pedagogies of Engagement in Science Learning

By Jose P. Mestre, professor of physics, University of Massachusetts, Amherst

Interest in adopting "pedagogies of engagement" in science teaching has sharply increased over the last decade. It turns out this is not a passing fad, but more a realization that teaching strategies that are consistent with research on learning must engage learners in constructing and making sense of their own knowledge. The contemporary view of learning, based on a large body of evidence from research studies, indicates that learners not only construct knowledge, but the knowledge they already possess filters any new knowledge that they are trying to learn. If new knowledge conflicts with previously constructed knowledge, the new knowledge will not make sense to the learner and may be constructed in a way that is incompatible with current scientific thought, or perhaps not useful for flexible application. Hence, sense making is central to the constructivist view of learning.

Constructivism has important implications for learning and instruction. First, construction of knowledge does not just happen in the classroom; rather, it is a lifelong, effortful process requiring significant mental engagement from the learner. Further, since resident knowledge filters the ability to construct new knowledge, it is important to keep in mind that the learner's mind is not a blank slate upon which new knowledge can be inscribed. Each learner comes into a classroom with a brain already wired by previous experiences. Depending on the existing connections, even the same concrete experiences are perceived differently by different learners.

From the perspective of instruction, teachers would be wise to probe students' previously constructed knowledge in order to make appropriate instructional choices so that students construct knowledge in ways that the teacher intends. Teachers then need to determine whether sufficient prior knowledge is available and evaluate whether this knowledge conflicts with the knowledge being taught. If conflict does exist, teachers should guide learners in reconstructing knowledge. The goal is to facilitate knowledge construction in learners that is compatible with scientific concepts so that learners will store the knowledge in memory in a form that is optimal for long-term recall and for application in problem-solving contexts (Etkina, Mestre, and O'Donnell, forthcoming). To ignore learners' preexisting knowledge makes it highly probable that the message intended by the teacher will not be the message understood by the student.

The most relevant, and perhaps obvious, instructional implication of constructivist epistemology is that teaching strategies that facilitate the construction of knowledge should be favored over those that do not. What, then, are pedagogical strategies that foster, encourage, and facilitate the construction of knowledge? Many pedagogies of engagement have emerged over the years with names such as cooperative or collaborative learning, active learning, case-based learning, and hands-on learning. All of these instructional strategies attempt to create an environment where students are actively thinking about and applying knowledge during the course of instruction, as opposed to passively listening to an instructor present the material in the textbook. This approach places more responsibility on the learner by expecting her or him to come to class prepared and ready to work at the difficult task of refining conceptual understanding and developing problem-solving skills.

The Emergence of Pedagogies of Engagement in Physics

An interesting series of events led to a sharp increase in the development and adoption of different forms of active learning in physics. Over twenty-five years ago, physicists and educators had begun to realize that students in physics classes, both in high school and college, had strong beliefs about how the physical world worked—beliefs that were in conflict with how physicists described the world. Physics education research at the time largely consisted of exploring how what came to be known as "misconceptions" interfered with the learning of physics concepts. Word about students' misconceptions had begun to trickle down to professors teaching introductory physics, but most paid little attention to the implications of this work for instruction; the view appeared to be that good professors were able to help students overcome these erroneous ideas with appropriate clarity and organization in the lectures.

Then, in the mid-1980s and early 1990s, David Hestenes and his collaborators (Halloun and Hestenes 1985; Hestenes, Wells, and Swackhamer 1992) developed and refined a simple test of basic misconceptions in mechanics that came to be known as the Force Concept Inventory (FCI). Figure 1 shows two questions from this test, which should make it evident that the FCI measures understanding of basic conceptual knowledge that physicists expect students to grasp after taking an introductory physics course. The FCI had a pronounced effect on dialogue among physicists about pedagogy, but the mechanism by which this dialogue began was somewhat amusing. Many who knew about the adverse effect of misconceptions on learning invited professors of introductory courses to administer the FCI to their classes at the end of the course. At first, not many took up this offer, stating that it was a waste of time to give such a basic exam to students following instruction since students would surely get all of those easy questions correct. As those few who actually administered the FCI quickly realized, students were finishing introductory courses and earning very good grades and yet had a very poor grasp of the concepts underlying the equations that they were adept at manipulating.

An anecdote that physics professor Eric Mazur at Harvard often tells about his personal experience with the FCI exemplifies the situation. Mazur tells that he was among those who thought that his students would surely know the answers to all of the questions in the FCI, but decided to try giving the exam to his class. He immediately knew that something was wrong when a student raised her hand and asked him whether he preferred that she answer the questions according to what he taught them, or according to what she really believed to be true. This incident was an epiphany for Mazur, who later developed and adopted an active learning approach that he calls "peer instruction" (Mazur 1997). What eventually became clear was that students taught via active engagement methods had much higher pre-post gains on the FCI than students taught by conventional lecture methods (Hake 1998), and furthermore, the charisma of the instructor had little to do with student gains on the FCI.

Features of Active Learning

There are several features commonly found in pedagogies of engagement:

- Students are actively engaged in constructing knowledge, often by working collaboratively on meaningful questions that are discussed within the context of concepts and procedures being covered in class, or of previously constructed knowledge. The questions do not have to be difficult, but care should be taken by the instructor so that they illustrate the meaning or application of major ideas in the course (e.g., questions that elicit possible misconceptions are very useful).
- Students voice the reasoning leading to their answers for evaluation by peers and by the instructor. Students thereby learn about constructing coherent arguments as well as evaluating arguments, and

teachers play the role of learning coach rather than dispensers of information.

- Class time is largely spent on refining conceptual understanding and on exploring procedures for applying conceptual knowledge across multiple relevant contexts. This helps students construct knowledge that is consistent with current scientific understanding and can be flexibly applied.
- Students rely less on the teacher as an authority figure, or as the "keeper of the knowledge," and take more responsibility for becoming self-sufficient learners. The focus of pedagogies of engagement is on the student, not the instructor.

Myths about Pedagogies of Engagement

I believe that several myths prevent pedagogies of engagement from being the norm in college science instruction:

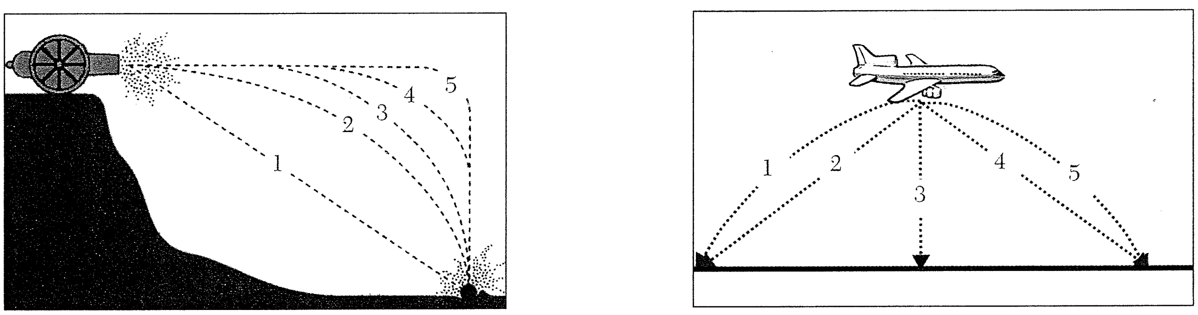
- *Myth 1: Pedagogies of engagement result in less content being covered and this is a disservice to students.* Part of this "myth" is grounded in fact—it is virtually guaranteed that the time needed to address students' difficulties during instruction means that less content will be covered, but just because active learning is less efficient for covering lots of content does not mean that we are doing students a disservice. Keep in mind that covering lots of content in class does not mean that students learn it, at least at the level and with the longevity that we desire. Active learning focuses on helping students understand somewhat less material but at a much deeper level, and this has more lasting effects than covering large amounts of content superficially.
- *Myth 2: Research on learning indicates that all lecturing is bad and to be avoided.* Not so. Although it is true that lecturing exclusively is not the most efficient way to help most students construct knowledge during class, there is nothing in the research literature that states that lecturing is useless for learning. In fact, there is research that hints at when lecturing is most effective. Bransford and Schwartz (1999) have shown that certain active learning activities prepare students to learn from lectures that follow those activities, even when little apparent progress is made in achieving the goals of the activity. These findings suggest that the activity primed students to learn from the subsequent lecture in a way that was not possible by lecturing alone.
- *Myth 3: I'm a great teacher because I get excellent teaching evaluations, and so there is no need to change my lecturing.* Unfortunately, teaching evaluations are not designed to measure how much students learn. Often they measure the charisma, showmanship, or popularity of the instructor. It would seem that instructors who consistently receive high teaching evaluations would be open to trying something new that might result in more learning in students, but my experience suggests the opposite.
- *Myth 4: Any activity that makes students active during class will result in more learning.* It is important to design classroom activities in ways that are optimal for the outcomes that the teacher desires, and this takes effort and experience. Having students discuss whether they liked yesterday's weather will not help them learn meteorological concepts.
- *Myth 5: Having students present their reasoning during class confuses other students and therefore is not worth the effort.* This is an argument that I often hear from many of my own students; they would rather that I just explain things because, they claim, listening to several erroneous arguments presented by other students causes them to remember wrong information. My counterargument to them is that equally important as constructing good arguments is judging the validity of arguments—a skill that scientists value highly.

Final Thoughts

While it is certainly true that much progress has been made in the last few decades on the nature of human learning and on effective teaching strategies, much remains to be determined about the definitive conditions that produce successful learners. However, in refuting the previously stated myths that all too often prevent pedagogies of engagement from being utilized in the classroom, it is clear that many of the commonly accepted views on how students learn should be challenged. Research has shown that through true pedagogies of engagement—actively engaging students in constructing knowledge with adequate classroom time for refining conceptual knowledge—we can create effective learning environments. ■

Figure 1:

Two questions from the Force Concept Inventory. All questions in this exam measure understanding of basic concepts covered in the typical introductory physics course.



FCI Question #12 (Cannon)
A ball is fired by a cannon from the top of a cliff as shown above. Which of the paths 1-5 would the cannon ball most closely follow?

FCI Question #23 (Airplane)
A bowling ball accidentally falls out of the cargo bay of an airliner as it flies in a horizontal direction.

As observed by a person standing on the ground and viewing the plane as in the above figure, which path would the bowling ball most closely follow after leaving the airplane?

(Artwork by Sanjay Rebello, used with permission.)

References

Bransford, J. D., and D. Schwartz. 1999. Rethinking transfer: A simple proposal with multiple implications. In *Review of Research in Education,* vol. 24, ed. A. Iran-Nejad and P. D. Pearson, 61-100. Washington, DC: American Educational Research Association.

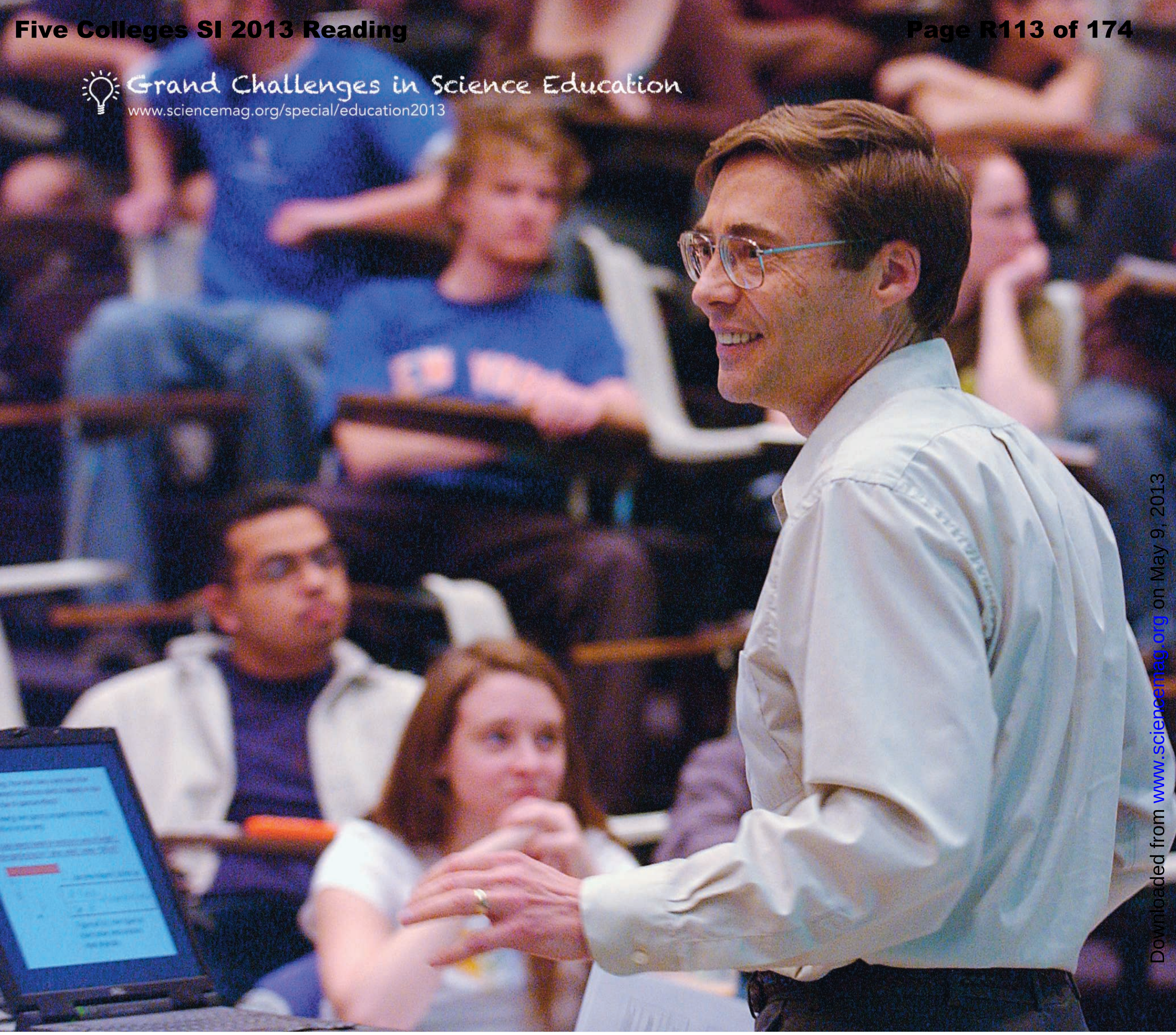
Etkina, E., J. Mestre, and A. O'Donnell. Forthcoming. The impact of the cognitive revolution on science learning and teaching. In *The cognitive revolution in educational psychology,* ed. J. M. Royer. Greenwich, CT: Information Age Publishing.

Hake, R. 1998. Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics* 66: 64-74.

Halloun, I. A., and D. Hestenes. 1985. Common sense concepts about motion. *American Journal of Physics* 53:1056-65.

Hestenes, D., M. Wells, and G. Swackhamer. 1992. Force concept inventory. *The Physics Teacher* 30 (3): 159-66.

Mazur, E. 1997. *Peer instruction: A user's manual.* Upper Saddle River, NJ: Prentice Hall.



GRAND CHALLENGE: UNDERGRADUATE TEACHING

Transformation Is Possible If a University Really Cares

The same attention to scientific detail that led to his Nobel Prize is helping Carl Wieman improve how undergraduates learn science

The way that most research universities across North America teach science to undergraduates is worse than ineffective, says Carl Wieman. It's unscientific.

A Nobel Prize–winning physicist turned science educator, Wieman doesn't understand why institutions of higher education would disregard decades of research showing the superiority of student-centered, active learning over the traditional 50-minute lecture. Using that outdated approach, he says, means universities are squandering talent at a time when U.S. higher education is being criticized for not turning out enough science-savvy graduates to keep the country competitive.

Wieman has spent the past 15 years applying the science of learning to how undergraduate science courses are taught. First at the University of Colorado, Boulder, (colorado.edu/sei) and, more recently, at the University of British Columbia (UBC), Vancouver, in Canada (cwsei.ubc.ca), Wieman and his colleagues have made impressive strides in changing how individual faculty members teach. Those changes, within individual courses, have translated into big improvements in student learning.

Those courses are offered by academic departments, which are his real target. Departments define the reward structure

CREDIT: CASEY A. CASS/UNIVERSITY OF COLORADO

Published by AAAS

Engaged instruction. Carl Wieman uses active learning tools to teach an undergraduate course at the University of Colorado in 2001.

for faculty members through their authority to hire, promote, and grant tenure, he says. So the best way to sustain improvements in teaching and learning is to get departments to buy into the need to change the courses that they offer. And that's begun to happen at UBC, one of Canada's elite universities.

Wieman's passion for the subject, combined with his stature as a Nobelist, has focused national attention on the high attrition rate among students who declare an interest in earning a degree in a STEM (science, technology, engineering, and mathematics) field. It's one of the biggest impediments to any effort to train more scientists and engineers. "I think Carl, more than anybody else, put a spotlight on the need to improve undergraduate education," says Subra Suresh, who last month stepped down as director of the National Science Foundation (NSF). "It wasn't a surprise to universities, but his work has highlighted the problem."

Colleagues also laud Wieman's rigorous approach to reform. "I have an incredible amount of respect for his deep commitment to the evidence," says Susan Singer, head of undergraduate education at NSF and a national leader in reforming undergraduate biology education. "Carl is someone who digs in and really wants to know."

Notwithstanding his success at Colorado and UBC, Wieman has made much less progress toward another of his goals: overturning an academic culture that values research over teaching. Working in the White House Office of Science and Technology Policy (OSTP) as associate director for science, Wieman was the de facto science education czar for the Obama administration. But his 20 months on the job taught him just how hard it is to change prevailing attitudes within U.S. higher education.

While at OSTP, Wieman floated the idea of requiring universities to collect and disseminate information on their teaching practices to remain eligible for federal research dollars. The policy would be a stick to get universities to pay more attention to teaching, he reasoned.

"There's an entire industry devoted to measuring how important my research is, with impact factors of papers and so on," Wieman says. "Yet, we don't even collect data on how I am teaching. It receives no attention. … If everything about teaching remains hidden, then universities can avoid having to devote anything more than minimal effort to doing it well. They can instead spend most of their time and money on research."

Wieman pushed the idea at numerous meetings with other government science officials and academic leaders. But they recoiled in horror at the prospect of what they viewed as another unfunded federal mandate. They prefer a 5-year effort begun last year by the 62-member Association of American Universities that aims to create a voluntary "framework" for improving teaching practices that institutions can adapt to their own situation and implement at their own pace.

"I'm very supportive of improving undergraduate STEM teaching," says Francis Collins, director of the National Institutes of Health (NIH), which spends more money on academic research than any other federal agency. "But this struck people as the wrong pathway by which to achieve the desired outcome, and not very fair."

> "There's an entire industry devoted to measuring how important my research is, with impact factors of papers and so on. Yet, we don't even collect data on how I am teaching."
>
> —Carl Wieman

As if taking on the nation's research establishment wasn't enough of a challenge, last June, Wieman received a sudden diagnosis of multiple myeloma. To deal with this health crisis, he abruptly resigned from his White House post and enrolled in a clinical trial at NIH using two experimental drugs. That treatment ended in January, and the 62-year-old Wieman says he's "happy and healthy."

Online
sciencemag.org
Podcast interview with author Jeffrey Mervis (http://scim.ag/pod_6130).

Wieman, who is leaving UBC but declined to say where he's going, has returned to the lecture circuit with an updated version of his standard talk, entitled "Taking a Scientific Approach to Science and Engineering Education." He's definitely not cowed by the prospect of taking a long, hard road toward his goal. In fact, his personal metric for any reform worth attempting is its ability "to generate significant opposition." Speaking at a session of the February annual meeting of AAAS (which publishes *Science*) in Boston, Wieman said that transforming undergraduate teaching "passes that litmus test."

Giving reform a chance

Wieman's personality and upbringing seem well-suited to a grand challenge like remaking undergraduate science education. Before deciding on a scientific career, he embraced a succession of passions, including chess and tennis, which for a time were all-consuming. "Monomaniacal pretty much describes me," Wieman confessed during a 2007 interview with the Nobel committee. "My view of everything is that you become good at something by focusing and working hard at it." Eventually, he recalls, "science [became] such an activity."

That doggedness served him well in pursuing his Nobel Prize–winning research. In 1925, Albert Einstein, building on the work of Indian physicist Satyendra Nath Bose, deduced that cooling a gas of certain atoms should make all the atoms suddenly flop into the same lowest energy quantum wave. Such a macroscopic matter wave is known as a Bose-Einstein condensate. Some 70 years later, Wieman and Eric Cornell of JILA, a lab run jointly by the National Institute of Standards and Technology and the University of Colorado, Boulder, achieved one by employing magnets and lasers to cool rubidium-87 atoms to within a millionth of a degree of absolute zero. In 2001, the two physicists shared the Nobel Prize in physics with Wolfgang Ketterle of the Massachusetts Institute of Technology in Cambridge, who achieved a similar result with sodium-23 atoms.

Wieman and Cornell at least had the advantage of knowing what a Bose-Einstein condensate would look like before they created one. In contrast, many faculty members might not recognize high-quality, student-centered learning because they may never have experienced it. Wieman admits that many notable scientists have thrived on a diet of traditional teaching practices and that the current rewards system at most universities gives faculty members little reason to try something different.

"I'm certainly not one to dismiss the importance of research," Wieman says. "But people need to recognize how totally dominant the reward system is. There are a lot of faculty who feel, completely appropriately, that 'I could spend more time improving my teaching, but that's not what I'm supposed to be doing.' So you have to figure out a way for them to be able to improve their teaching without making a big sacrifice in their research activities."

Wieman embarked on his quest to improve undergraduate education after pondering his own career as a professor and educator. And like the scientist that he is, he began by asking himself some basic questions. Why, he wondered, did students in his introductory courses do so poorly, and even regress, after he delivered lectures covering what they needed to know? Why couldn't he identify at the outset which graduate students were most likely to succeed? And why did most of them become productive scientists after a few years in his lab?

Digging into the literature on teaching and learning yielded some insights. His graduate students had learned to think like scientists, he realized, by doing real science under the supervision of a world-class scientist. Developing expertise, he came to understand, is a slow and arduous process marked by repeated failures.

> "I know that Carl is skeptical universities will do it on their own. But I have yet to be convinced that they won't."
>
> —Francis Collins, NIH director

"The apprentice model works pretty well in graduate school because the faculty member can see if the student is learning how to build a laser system, or write a paper, or give a professional talk," says University of Colorado, Boulder, physicist and education researcher Noah Finkelstein, who has worked closely with Wieman and now directs the university's newly formed Center for STEM Learning. "Those are things we actually want them to do. We give them feedback along the way, and we take in feedback from them and adjust our mentoring. But that system is just too costly at the undergraduate level."

Instead, faculty members must interact with hundreds of students in a large hall. Most choose to do that via a lecture. But research has shown that most students cling to their misconceptions even after sitting through a brilliant lecture.

What works better than lectures and homework problems, according to numerous studies, is having students work in small teams with instructors who can help them apply those basic concepts to real-life situations. But what's the best way to implement active, student-centered learning? The answer, Wieman decided, lay in melding it with the concept of deliberate practice.

That idea, developed by psychologist K. Anders Ericsson of Florida State University in Tallahassee, treats the brain as a muscle that must be exercised to perform at its peak. It's how a novice becomes an expert, whether in music, sports, or science. "We have learned that complex expertise is a matter not of filling up an existing brain with knowledge, but of brain development," Wieman says.

Deliberate practice, Wieman wrote in the fall 2012 issue of *Issues in Science and Technology*, "involves the learner solving a set of tasks or problems that are challenging but doable and that involve explicitly practicing the appropriate expert thinking and performance." The teacher, or coach, offers appropriate incentives to encourage students to master the necessary skills, as well as continuous feedback to help them remain on task. As with any sport, he notes, "[t]housands of hours of deliberate practice are typically required to reach an elite level of performance."

The two concepts created an intellectual framework around which to transform undergraduate science. "Just as we have physics principles, here are the principles that work, and they are consistent with what others had done," Wieman says. "It also allows you to go into disciplines where there hadn't been much work done, like oceanography, and make some generalizations. It's very much like science itself."

In a 2011 paper in *Science,* Wieman and his colleagues describe the power of active learning and deliberate practice. The instructor for one section of an introductory physics class for engineers at UBC used these principles, while the other instructor delivered the normal lectures. The first group of students scored more than twice as high on a multiple-choice test of the material covered than did those in the control group.

"The results were so dramatic from this relatively modest experiment that the entire [physics] department had an epiphany," remarks Simon Peacock, UBC's dean of sciences. "It sent them a clear message: Wow, we can actually teach better."

Wieman says that active learning and deliberate practice is now the norm in 99 UBC courses enrolling 31,200 students. Many are introductory courses taken by freshmen and sophomores who are still uncertain of their major field of study. "We have substantially changed more than half of the math and science courses a UBC student in the college of science will take in their first 2 years," Wieman says, citing results from a recent survey of how faculty members have changed their teaching practices since the Carl Wieman Science Education Initiative was launched in 2007.

"We've hit it out of the park with earth and ocean sciences," one of seven departments that are part of the university-funded initiative, Peacock says. "I will declare them to be a success."

Wieman believes that deliberate practice can also help students in primary and secondary school who, for whatever reason, are ill-prepared for success in STEM subjects. His efforts have helped resolve "a huge controversy," says NSF's Singer, over whether the vast majority of students are capable of doing high-level math and science.

"Having Carl stand up and say we should stop doing STEM talent selection and start doing STEM talent development completely changes the nature of the conversation," says Singer, on leave from her post as a biology professor at Carleton College in Northfield, Minnesota. "It's really a question of how you structure the learning environment. And his work has shown that active learning strategies are more effective."

From my way to the right way

What does it take to transform an undergraduate science course? Wieman's approach relies heavily on a cadre of science teaching and learning fellows, who are typically postdocs. At its height, the Colorado initiative employed a dozen such fellows; at UBC, the number peaked at nearly two dozen.

The fellows are trained in the many steps needed to transform a university lecture course—steps that faculty members are unlikely to take on their own, either out of ignorance or because they simply don't have the time to do what's needed. Katherine Perkins, who directs both the science education initiative at Colorado and the related PhET project (phet.colorado.edu), which has created thousands of research-based simulations of physical phenomena, calls the teaching and learning fellows "engines of change."

Meeting for the first time with a faculty member, a fellow might start by asking what the faculty member wants students to know how to do at the end of the course. That's a more useful metric than asking what a student "should understand," explains Beth Simon, director of the Center for Teaching Development at the University of California, San Diego, who spent the 2007 to 2008 academic year at UBC as a fellow in the computer science department before returning to UCSD.

Once the faculty member articulates the real goals of the course, those skills are converted into learning objectives. The next step is to write up multiple-choice questions aimed at helping students achieve each learning objective. The so-called clicker questions (the name comes from the electronic device that students use to record their answers) usually focus on common student misconceptions about the concepts.

The questions become the basic curriculum for the course. But getting from skills to clicker questions can be difficult. Simon figured that the final exam would provide a useful guide to what students were expected to learn. Instead, instructors would admit that they didn't really know what concepts some test questions were meant to measure, she says, and that other questions covered concepts not central to the course.

Most courses come with only a three- or four-sentence description in the syllabus. That brevity gives whoever is teaching the course a lot of leeway. Some faculty members have been teaching the same course for years, Simon says, and for them, "learning outcomes were a nonstarter. 'I teach 101 my way,' they would say." In contrast, some courses are "owned" by the department and a consensus exists on what students are expected to know regardless of who is teaching the course.

A transformed course typically begins not with a lecture but with a clicker question. Students gather in small groups to discuss it, and a fellow assigned to the course circulates through the classroom to guide the inquiry process. Once the students have punched in their answers, the faculty member might offer a microlecture aimed at correcting their mistakes and filling in gaps in their knowledge. Once the concept is clear, the class moves on to the next clicker question.

> "Having Carl stand up and say we should stop doing STEM talent selection and start doing STEM talent development completely changes the nature of the conversation."
>
> —Susan Singer, head of NSF undergraduate education

Students taking transformed courses are usually more active than in a typical lecture class. Faculty members need to remind students regularly why they will not be lecturing, Simon says, as well as explain the importance of peer instruction. To get the most from the class time, students are assigned outside reading and turn in homework that measures their understanding of the material.

Some students are uncomfortable with this approach—even if it's more effective. "I remember getting an evaluation from one [UCSD] student who had just finished my course," says Simon, a pioneer in the use of peer instruction within her field. "I loved it. It read, 'I just wish she'd have lectured. Instead, I had to learn the material myself.'"

The increased student engagement in a transformed course is music to the ears of the average faculty member. "Most faculty want their students to learn more," says Perkins, whom Wieman hired in 2003 as one of the initiative's first teaching fellows. "They look at the final exam, sigh, and say, 'Why did only 60% get that question right?' " Simon adds, "If they can have more fun, they will choose to use these methods."

A department should plan on spending about 5% of its budget for 5 years to transform its courses, Wieman says. Lesser amounts are required to sustain progress, he adds, although new faculty members must be trained and existing faculty members need ongoing support and, occasionally, a sympathetic ear. At Colorado, for example, departments competed for grants of roughly $600,000 to $800,000 each. UBC's $10 million commitment to the initiative allowed Wieman to double the size of departmental awards, and a more recent $2 million donation from David Cheriton, a professor of computer science at Stanford University, is fueling reform within the math and computer science department.

Wieman's campaign to transform departments isn't the only game in town. Finkelstein's new center at Colorado, funded by an NSF planning grant, is supposed to serve as a focal point for some 75 STEM-related activities on campus. And Colorado's Perkins hopes that NSF will put up several million dollars for a Web site to help faculty members use the PhET simulations that she and others have created and to study their impact on teaching and learning.

But money remains tight. Wieman says he can't afford to conduct the rigorous, outside assessments that normally accompany NSF-funded reforms because he feels that institutional funds should redound to the benefit of the institution. However, the dearth of peer-reviewed publications has led some scientists to question what Wieman's Colorado and UBC initiatives have accomplished.

"When people ask what we've done," Finkelstein says, "and I say we've shifted institutional identity and culture, half the time their response is, 'Wow, that's terrific.' But the other half say, 'So all you've done is talk?' "

Wieman himself offers a frank answer when asked whether he expects the UBC reforms to stick. "That's why you do research," he says. "This was a one-time intervention. And people have a right to wonder what will happen next.

"I'm more optimistic than I was a year ago," he adds, "because people who we thought weren't interested are now saying, 'Look, I made this change and I'm thinking of doing more.' But I won't give you good odds that they will still be doing it in 10 years."

Carrot or stick?

In 2010, Wieman decided to come to Washington for the chance to influence undergraduate science education on a national scale. "My top priority at OSTP was to improve undergraduate education," he says. "We know what to do that will help students learn more and be more successful and how to get a broader group of students doing it."

While there, Wieman came up with his simple, market-driven first step: Require universities to compile and release data on their teaching methods as a condition for receiving federal research funds. As students began using the data released by universities to help choose a college, he reasoned, universities would feel compelled to improve their teaching practices in order to attract the best applicants. "If an agency were to require every grantee to provide this information," he says, "then the next year teaching would look completely different because somebody is looking at it."

> "There are a lot of faculty who feel, completely appropriately, that 'I could spend more time improving my teaching, but that's not what I'm supposed to be doing.'"
>
> —Carl Wieman

Wieman promoted the idea tirelessly in meetings with his government colleagues as well as the presidents of several leading research universities, seeing it as a painless way to propel reform. But they pushed back hard. It's hard to define particular teaching practices, they told Wieman. Self-reported data are unreliable, they added, and collecting such data would be a burden. Last April, the presidents of several prominent universities even wrote a letter to then–White House Chief of Staff Jacob Lew in an attempt to head off Wieman's proposal.

A few months later, Wieman was gone. But he hasn't changed his mind one iota, and he says that none of the community's objections are valid.

For starters, he says, colleagues at UBC and Colorado have created a questionnaire that collects such data and requires only a few hours of effort by an entire department—"a tiny amount compared to what is spent in a single faculty meeting," he snickers. Universities have no incentive to game the system, he adds, because students would soon expose any institution that had submitted bogus information. And he scoffs at the idea that tracking a fundamental purpose of a university could be regarded as a "burden."

Part of their objections, he speculates, is that the data could prove embarrassing. "Educational transparency is a threat to their status," he argues. "Maybe it won't make them look so good."

NIH's Collins says that's not the reason he prefers a voluntary approach. "I know that Carl is skeptical universities will do it on their own," he says. "But I have yet to be convinced that they won't. I don't know that all universities will want to participate. But I think there will be some who would say, 'Yeah, we believe in this. It's the right thing to do.' "

Government officials and university leaders typically defend the value of federally funded research by citing its role as an engine of economic growth. In the case of biomedical research, they also note its potential to save lives. But Wieman doesn't think those arguments really address the growing clamor from the public and politicians for universities to show that an undergraduate education is worth the rising cost of tuition. That skepticism, he says, has also fueled a decadelong assault by many state legislatures on their flagship public universities.

A more effective response, Wieman says, would be for university presidents to emphasize how research can lead to better teaching. "I think the solution is to show that you can really use that research expertise to improve education," he says. "Deliberate practice and other approaches is calling on, and demanding of, the research expertise embodied by that faculty."

"If you pitch that message," he continues, "then suddenly it becomes clear how having a great research university translates into better education for students in my state. Right now it's not worth the investment, because it's not happening. But it could."

–JEFFREY MERVIS

Science faculty's subtle gender biases favor male students

Corinne A. Moss-Racusin[a,b], John F. Dovidio[b], Victoria L. Brescoll[c], Mark J. Graham[a,d], and Jo Handelsman[a,1]

[a]Department of Molecular, Cellular and Developmental Biology, [b]Department of Psychology, [c]School of Management, and [d]Department of Psychiatry, Yale University, New Haven, CT 06520

Edited* by Shirley Tilghman, Princeton University, Princeton, NJ, and approved August 21, 2012 (received for review July 2, 2012)

Despite efforts to recruit and retain more women, a stark gender disparity persists within academic science. Abundant research has demonstrated gender bias in many demographic groups, but has yet to experimentally investigate whether science faculty exhibit a bias against female students that could contribute to the gender disparity in academic science. In a randomized double-blind study (*n* = 127), science faculty from research-intensive universities rated the application materials of a student—who was randomly assigned either a male or female name—for a laboratory manager position. Faculty participants rated the male applicant as significantly more competent and hireable than the (identical) female applicant. These participants also selected a higher starting salary and offered more career mentoring to the male applicant. The gender of the faculty participants did not affect responses, such that female and male faculty were equally likely to exhibit bias against the female student. Mediation analyses indicated that the female student was less likely to be hired because she was viewed as less competent. We also assessed faculty participants' preexisting subtle bias against women using a standard instrument and found that preexisting subtle bias against women played a moderating role, such that subtle bias against women was associated with less support for the female student, but was unrelated to reactions to the male student. These results suggest that interventions addressing faculty gender bias might advance the goal of increasing the participation of women in science.

diversity | lifestyle choices | science education | science workforce

A 2012 report from the President's Council of Advisors on Science and Technology indicates that training scientists and engineers at current rates will result in a deficit of 1,000,000 workers to meet United States workforce demands over the next decade (1). To help close this formidable gap, the report calls for the increased training and retention of women, who are starkly underrepresented within many fields of science, especially among the professoriate (2–4). Although the proportion of science degrees granted to women has increased (5), there is a persistent disparity between the number of women receiving PhDs and those hired as junior faculty (1–4). This gap suggests that the problem will not resolve itself solely by more generations of women moving through the academic pipeline but that instead, women's advancement within academic science may be actively impeded.

With evidence suggesting that biological sex differences in inherent aptitude for math and science are small or nonexistent (6–8), the efforts of many researchers and academic leaders to identify causes of the science gender disparity have focused instead on the life choices that may compete with women's pursuit of the most demanding positions. Some research suggests that these lifestyle choices (whether free or constrained) likely contribute to the gender imbalance (9–11), but because the majority of these studies are correlational, whether lifestyle factors are solely or primarily responsible remains unclear. Still, some researchers have argued that women's preference for nonscience disciplines and their tendency to take on a disproportionate amount of child- and family-care are the primary causes of the gender disparity in science (9–11), and that it "is not caused by discrimination in these domains" (10). This assertion has received substantial attention and generated significant debate among the scientific community, leading some to conclude that gender discrimination indeed does not exist nor contribute to the gender disparity within academic science (e.g., refs. 12 and 13).

Despite this controversy, experimental research testing for the presence and magnitude of gender discrimination in the biological and physical sciences has yet to be conducted. Although acknowledging that various lifestyle choices likely contribute to the gender imbalance in science (9–11), the present research is unique in investigating whether faculty gender bias exists within academic biological and physical sciences, and whether it might exert an independent effect on the gender disparity as students progress through the pipeline to careers in science. Specifically, the present experiment examined whether, given an equally qualified male and female student, science faculty members would show preferential evaluation and treatment of the male student to work in their laboratory. Although the correlational and related laboratory studies discussed below suggest that such bias is likely (contrary to previous arguments) (9–11), we know of no previous experiments that have tested for faculty bias against female students within academic science.

If faculty express gender biases, we are not suggesting that these biases are intentional or stem from a conscious desire to impede the progress of women in science. Past studies indicate that people's behavior is shaped by implicit or unintended biases, stemming from repeated exposure to pervasive cultural stereotypes (14) that portray women as less competent but simultaneously emphasize their warmth and likeability compared with men (15). Despite significant decreases in overt sexism over the last few decades (particularly among highly educated people) (16), these subtle gender biases are often still held by even the most egalitarian individuals (17), and are exhibited by both men and women (18). Given this body of work, we expected that female faculty would be just as likely as male faculty to express an unintended bias against female undergraduate science students. The fact that these prevalent biases often remain undetected highlights the need for an experimental investigation to determine whether they may be present within academic science and, if so, raise awareness of their potential impact.

Whether these gender biases operate in academic sciences remains an open question. On the one hand, although considerable research demonstrates gender bias in a variety of other domains (19–23), science faculty members may not exhibit this

Author contributions: C.A.M.-R., J.F.D., V.L.B., M.J.G., and J.H. designed research; C.A.M.-R. performed research; C.A.M.-R. analyzed data; and C.A.M.-R., J.F.D., V.L.B., M.J.G., and J.H. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: jo.handelsman@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1211286109/-/DCSupplemental.

bias because they have been rigorously trained to be objective. On the other hand, research demonstrates that people who value their objectivity and fairness are paradoxically particularly likely to fall prey to biases, in part because they are not on guard against subtle bias (24, 25). Thus, by investigating whether science faculty exhibit a bias that could contribute to the gender disparity within the fields of science, technology, engineering, and mathematics (in which objectivity is emphasized), the current study addressed critical theoretical and practical gaps in that it provided an experimental test of faculty discrimination against female students within academic science.

A number of lines of research suggest that such discrimination is likely. Science is robustly male gender-typed (26, 27), resources are inequitably distributed among men and women in many academic science settings (28), some undergraduate women perceive unequal treatment of the genders within science fields (29), and nonexperimental evidence suggests that gender bias is present in other fields (19). Some experimental evidence suggests that even though evaluators report liking women more than men (15), they judge women as less competent than men even when they have identical backgrounds (20). However, these studies used undergraduate students as participants (rather than experienced faculty members), and focused on performance domains outside of academic science, such as completing perceptual tasks (21), writing nonscience articles (22), and being evaluated for a corporate managerial position (23).

Thus, whether aspiring women scientists encounter discrimination from faculty members remains unknown. The formative predoctoral years are a critical window, because students' experiences at this juncture shape both their beliefs about their own abilities and subsequent persistence in science (30, 31). Therefore, we selected this career stage as the focus of the present study because it represents an opportunity to address issues that manifest immediately and also resurface much later, potentially contributing to the persistent faculty gender disparity (32, 33).

Current Study

In addition to determining whether faculty expressed a bias against female students, we also sought to identify the processes contributing to this bias. To do so, we investigated whether faculty members' perceptions of student competence would help to explain why they would be less likely to hire a female (relative to an identical male) student for a laboratory manager position. Additionally, we examined the role of faculty members' preexisting subtle bias against women. We reasoned that pervasive cultural messages regarding women's lack of competence in science could lead faculty members to hold gender-biased attitudes that might subtly affect their support for female (but not male) science students. These generalized, subtly biased attitudes toward women could impel faculty to judge equivalent students differently as a function of their gender.

The present study sought to test for differences in faculty perceptions and treatment of equally qualified men and women pursuing careers in science and, if such a bias were discovered, reveal its mechanisms and consequences within academic science. We focused on hiring for a laboratory manager position as the primary dependent variable of interest because it functions as a professional launching pad for subsequent opportunities. As secondary measures, which are related to hiring, we assessed: (*i*) perceived student competence; (*ii*) salary offers, which reflect the extent to which a student is valued for these competitive positions; and (*iii*) the extent to which the student was viewed as deserving of faculty mentoring.

Our hypotheses were that: Science faculty's perceptions and treatment of students would reveal a gender bias favoring male students in perceptions of competence and hireability, salary conferral, and willingness to mentor (hypothesis A); Faculty gender would not influence this gender bias (hypothesis B); Hiring discrimination against the female student would be mediated (i.e., explained) by faculty perceptions that a female student is less competent than an identical male student (hypothesis C); and Participants' preexisting subtle bias against women would moderate (i.e., impact) results, such that subtle bias against women would be negatively related to evaluations of the female student, but unrelated to evaluations of the male student (hypothesis D).

Results

A broad, nationwide sample of biology, chemistry, and physics professors (n = 127) evaluated the application materials of an undergraduate science student who had ostensibly applied for a science laboratory manager position. All participants received the same materials, which were randomly assigned either the name of a male (n = 63) or a female (n = 64) student; student gender was thus the only variable that differed between conditions. Using previously validated scales, participants rated the student's competence and hireability, as well as the amount of salary and amount of mentoring they would offer the student. Faculty participants believed that their feedback would be shared with the student they had rated (see *Materials and Methods* for details).

Student Gender Differences. The competence, hireability, salary conferral, and mentoring scales were each submitted to a two (student gender; male, female) × two (faculty gender; male, female) between-subjects ANOVA. In each case, the effect of student gender was significant (all $P < 0.01$), whereas the effect of faculty participant gender and their interaction was not (all $P > 0.19$). Tests of simple effects (all $d > 0.60$) indicated that faculty participants viewed the female student as less competent [$t(125) = 3.89$, $P < 0.001$] and less hireable [$t(125) = 4.22$, $P < 0.001$] than the identical male student (Fig. 1 and Table 1). Faculty participants also offered less career mentoring to the female student than to the male student [$t(125) = 3.77$, $P < 0.001$]. The mean starting salary offered the female student, $26,507.94, was significantly lower than that of $30,238.10 to the male student [$t(124) = 3.42$, $P < 0.01$] (Fig. 2). These results support hypothesis A.

In support of hypothesis B, faculty gender did not affect bias (Table 1). Tests of simple effects (all $d < 0.33$) indicated that female faculty participants did not rate the female student as more competent [$t(62) = 0.06$, $P = 0.95$] or hireable [$t(62) = 0.41$, $P = 0.69$] than did male faculty. Female faculty also did not offer more mentoring [$t(62) = 0.29$, $P = 0.77$] or a higher salary [$t(61) = 1.14$, $P = 0.26$] to the female student than did their male

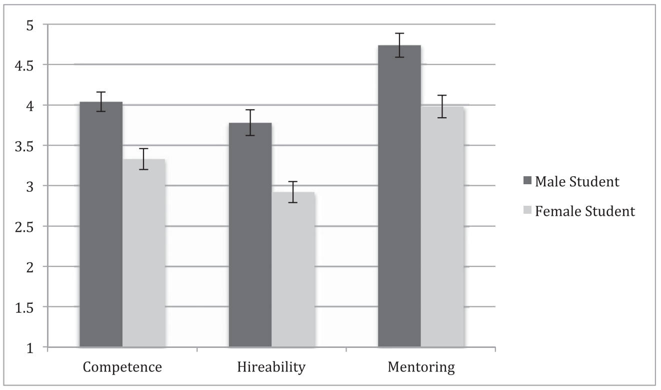


Fig. 1. Competence, hireability, and mentoring by student gender condition (collapsed across faculty gender). All student gender differences are significant ($P < 0.001$). Scales range from 1 to 7, with higher numbers reflecting a greater extent of each variable. Error bars represent SEs. $n_{\text{male student condition}} = 63$, $n_{\text{female student condition}} = 64$.

Table 1. Means for student competence, hireability, mentoring and salary conferral by student gender condition and faculty gender

	Male target student				Female target student				
	Male faculty		Female faculty		Male faculty		Female faculty		
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	*d*
Competence	4.01$_a$	(0.92)	4.1$_a$	(1.19)	3.33$_b$	(1.07)	3.32$_b$	(1.10)	0.71
Hireability	3.74$_a$	(1.24)	3.92$_a$	(1.27)	2.96$_b$	(1.13)	2.84$_b$	(0.84)	0.75
Mentoring	4.74$_a$	(1.11)	4.73$_a$	(1.31)	4.00$_b$	(1.21)	3.91$_b$	(0.91)	0.67
Salary	30,520.83$_a$	(5,764.86)	29,333.33$_a$	(4,952.15)	27,111,11$_b$	(6,948.58)	25,000.00$_b$	(7,965.56)	0.60

Scales for competence, hireability, and mentoring range from 1 to 7, with higher numbers reflecting a greater extent of each variable. The scale for salary conferral ranges from \$15,000 to \$50,000. Means with different subscripts within each row differ significantly ($P < 0.05$). Effect sizes (Cohen's *d*) represent target student gender differences (no faculty gender differences were significant, all $P > 0.14$). Positive effect sizes favor male students. Conventional small, medium, and large effect sizes for *d* are 0.20, 0.50, and 0.80, respectively (51). $n_{\text{male student condition}} = 63$, $n_{\text{female student condition}} = 64$. $^{***}P < 0.001$.

colleagues. In addition, faculty participants' scientific field, age, and tenure status had no effect (all $P > 0.53$). Thus, the bias appears pervasive among faculty and is not limited to a certain demographic subgroup.

Mediation and Moderation Analyses. Thus far, we have considered the results for competence, hireability, salary conferral, and mentoring separately to demonstrate the converging results across these individual measures. However, composite indices of measures that converge on an underlying construct are more statistically reliable, stable, and resistant to error than are each of the individual items (e.g., refs. 34 and 35). Consistent with this logic, the established approach to measuring the broad concept of target competence typically used in this type of gender bias research is to standardize and average the competence scale items and the salary conferral variable to create one composite competence index, and to use this stable convergent measure for all analyses (e.g., refs. 36 and 37). Because this approach obscures mean salary differences between targets, we chose to present salary as a distinct dependent variable up to this point, to enable a direct test of the potential discrepancy in salary offered to the male and female student targets. However, to rigorously examine the processes underscoring faculty gender bias, we reverted to standard practices at this point by averaging the standardized salary variable with the competence scale items to create a robust composite competence variable ($\alpha = 0.86$). This composite competence variable was used in all subsequent mediation and moderation analyses.

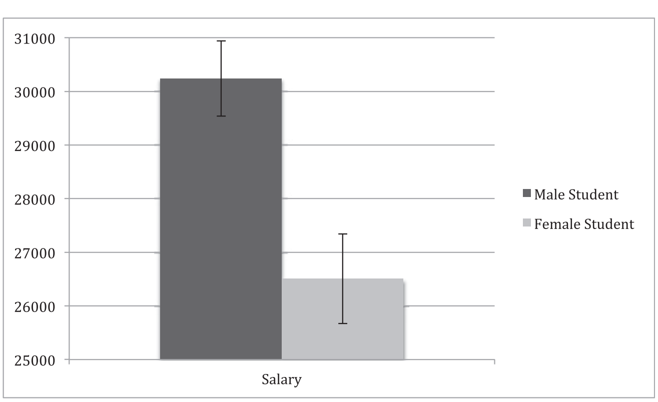


Fig. 2. Salary conferral by student gender condition (collapsed across faculty gender). The student gender difference is significant ($P < 0.01$). The scale ranges from \$15,000 to \$50,000. Error bars represent SEs. $n_{\text{male student condition}} = 63$, $n_{\text{female student condition}} = 64$.

Evidence emerged for hypothesis C, the predicted mediation (i.e., causal path; see *SI Materials and Methods: Additional Analyses* for more information on mediation and the results of additional mediation analyses). The initially significant impact of student gender on hireability ($\beta = -0.35$, $P < 0.001$) was reduced in magnitude and dropped to nonsignificance ($\beta = -0.10$, $P = 0.13$) after accounting for the impact of student composite competence (which was a strong predictor, $\beta = 0.69$, $P < 0.001$), Sobel's $Z = 3.94$, $P < 0.001$ (Fig. 3). This pattern of results provides evidence for full mediation, indicating that the female student was less likely to be hired than the identical male because she was viewed as less competent overall.

We also conducted moderation analysis (i.e., testing for factors that could amplify or attenuate the demonstrated effect) to determine the impact of faculty participants' preexisting subtle bias against women on faculty participants' perceptions and treatment of male and female science students (see *SI Materials and Methods: Additional Analyses* for more information on and the results of additional moderation analyses). For this purpose, we administered the Modern Sexism Scale (38), a well-validated instrument frequently used for this purpose (*SI Materials and Methods*). Consistent with our intentions, this scale measures unintentional negativity toward women, as contrasted with a more blatant form of conscious hostility toward women.

Results of multiple regression analyses indicated that participants' preexisting subtle bias against women significantly interacted with student gender to predict perceptions of student composite competence ($\beta = -0.39$, $P < 0.01$), hireability ($\beta = -0.31$, $P < 0.05$), and mentoring ($\beta = -0.55$, $P < 0.001$). To interpret these significant interactions, we examined the simple effects separately by student gender. Results revealed that the more preexisting subtle bias participants exhibited against women, the less composite competence ($\beta = -0.36$, $P < 0.01$) and hireability ($\beta = -0.39$, $P < 0.01$) they perceived in the female student, and the less mentoring ($\beta = -0.53$, $P < 0.001$) they were willing to offer her. In contrast, faculty participants' levels of preexisting subtle bias against women were unrelated to the perceptions of the male student's composite competence ($\beta = 0.16$, $P = 0.22$) and hireability ($\beta = 0.07$, $P = 0.59$), and the amount of mentoring ($\beta = 0.22$, $P = 0.09$) they were willing to offer him. [Although this effect is marginally significant, its direction suggests that faculty participants' preexisting subtle bias against women may actually have made them more inclined to mentor the male student relative to the female student (although this effect should be interpreted with caution because of its marginal significance).] Thus, it appears that faculty participants' preexisting subtle gender bias undermined support for the female student but was unrelated to perceptions and treatment of the male student. These findings support hypothesis D.

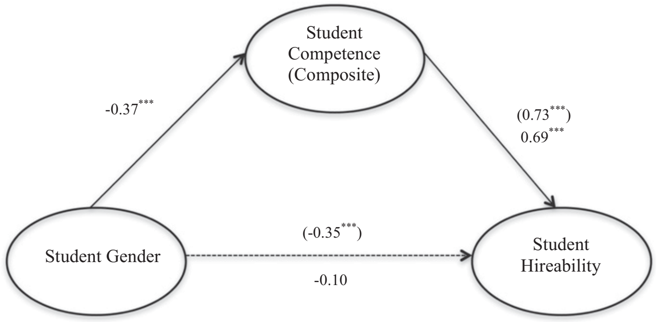


Fig. 3. Student gender difference hiring mediation. Values are standardized regression coefficients. The value in parentheses reflects a bivariate analysis. The dashed line represents the mediated path. The composite student competence variable consists of the averaged standardized salary variable and the competence scale items. Student gender is coded such that male = 0, female = 1. $n_{\text{male student condition}} = 63$, $n_{\text{female student condition}} = 64$. $^{***}P < 0.001$.

Finally, using a previously validated scale, we also measured how much faculty participants liked the student (see *SI Materials and Methods*). In keeping with a large body of literature (15), faculty participants reported liking the female (mean = 4.35, SD = 0.93) more than the male student [(mean = 3.91, SD = 0.1.08), $t(125) = -2.44$, $P < 0.05$]. However, consistent with this previous literature, liking the female student more than the male student did not translate into positive perceptions of her composite competence or material outcomes in the form of a job offer, an equitable salary, or valuable career mentoring. Moreover, only composite competence (and not likeability) helped to explain why the female student was less likely to be hired; in mediation analyses, student gender condition ($\beta = -0.48$, $P < 0.001$) remained a strong predictor of hireability along with likeability ($\beta = 0.60$, $P < 0.001$). These findings underscore the point that faculty participants did not exhibit outright hostility or dislike toward female students, but were instead affected by pervasive gender stereotypes, unintentionally downgrading the competence, hireability, salary, and mentoring of a female student compared with an identical male.

Discussion

The present study is unique in investigating subtle gender bias on the part of faculty in the biological and physical sciences. It therefore informs the debate on possible causes of the gender disparity in academic science by providing unique experimental evidence that science faculty of both genders exhibit bias against female undergraduates. As a controlled experiment, it fills a critical gap in the existing literature, which consisted only of experiments in other domains (with undergraduate students as participants) and correlational data that could not conclusively rule out the influence of other variables.

Our results revealed that both male and female faculty judged a female student to be less competent and less worthy of being hired than an identical male student, and also offered her a smaller starting salary and less career mentoring. Although the differences in ratings may be perceived as modest, the effect sizes were all moderate to large ($d = 0.60$–0.75). Thus, the current results suggest that subtle gender bias is important to address because it could translate into large real-world disadvantages in the judgment and treatment of female science students (39). Moreover, our mediation findings shed light on the processes responsible for this bias, suggesting that the female student was less likely to be hired than the male student because she was perceived as less competent. Additionally, moderation results indicated that faculty participants' preexisting subtle bias against women undermined their perceptions and treatment of the female (but not the male) student, further suggesting that chronic subtle biases may harm women within academic science. Use of a randomized controlled design and established practices from audit study methodology support the ecological validity and educational implications of our findings (*SI Materials and Methods*).

It is noteworthy that female faculty members were just as likely as their male colleagues to favor the male student. The fact that faculty members' bias was independent of their gender, scientific discipline, age, and tenure status suggests that it is likely unintentional, generated from widespread cultural stereotypes rather than a conscious intention to harm women (17). Additionally, the fact that faculty participants reported liking the female more than the male student further underscores the point that our results likely do not reflect faculty members' overt hostility toward women. Instead, despite expressing warmth toward emerging female scientists, faculty members of both genders appear to be affected by enduring cultural stereotypes about women's lack of science competence that translate into biases in student evaluation and mentoring.

Our careful selection of expert participants revealed gender discrimination among existing science faculty members who interact with students on a regular basis (*SI Materials and Methods: Subjects and Recruitment Strategy*). This method allowed for a high degree of ecological validity and generalizability relative to an approach using nonexpert participants, such as other undergraduates or lay people unfamiliar with laboratory manager job requirements and academic science mentoring (i.e., the participants in much psychological research on gender discrimination). The results presented here reinforce those of Stenpries, Anders, and Ritzke (40), the only other experiment we know of that recruited faculty participants. Because this previous experiment also indicated bias within academic science, its results raised serious concerns about the potential for faculty bias within the biological and physical sciences, casting further doubt on assertions (based on correlational data) that such biases do not exist (9–11). In the Steinpreis et al. experiment, psychologists were more likely to hire a psychology faculty job applicant when the applicant's curriculum vitae was assigned a male (rather than female) name (40). This previous work invited a study that would extend the finding to faculty in the biological and physical sciences and to reactions to undergraduates, whose competence was not already fairly established by accomplishments associated with the advanced career status of the faculty target group of the previous study. By providing this unique investigation of faculty bias against female students in biological and physical sciences, the present study extends past work to a critical early career stage, and to fields where women's underrepresentation remains stark (2–4).

Indeed, our findings raise concerns about the extent to which negative predoctoral experiences may shape women's subsequent decisions about persistence and career specialization. Following conventions established in classic experimental studies to create enough ambiguity to leave room for potentially biased responses (20, 23), the student applicants in the present research were described as qualified to succeed in academic science (i.e., having coauthored a publication after obtaining 2 y of research experience), but not irrefutably excellent. As such, they represented a majority of aspiring scientists, and were precisely the type of students most affected by faculty judgments and mentoring (see *SI Materials and Methods* for more discussion). Our results raise the possibility that not only do such women encounter biased judgments of their competence and hireability, but also receive less faculty encouragement and financial rewards than identical male counterparts. Because most students depend on feedback from their environments to calibrate their own worth (41), faculty's assessments of students' competence likely contribute to students' self-efficacy and goal setting as scientists, which may influence decisions much later in their careers. Likewise, inasmuch as the advice and mentoring that students receive affect their ambitions and choices, it is significant that the faculty in this study were less inclined to mentor women than men. This finding raises the possibility that women may opt out of academic science careers in part because of diminished competence judgments, rewards, and mentoring received in the early years of the careers. In sum, the predoctoral years represent a window during which students' experiences of faculty bias or encouragement are particularly likely to shape their persistence in academic science (30–33). Thus, the present study not only fills an important gap in the research literature, but also has critical implications for pressing social and educational issues associated with the gender disparity in science.

If women's decisions to leave science fields when or before they reach the faculty level are influenced by unequal treatment by undergraduate advisors, then existing efforts to create more flexible work settings (42) or increase women's identification with science (27) may not fully alleviate a critical underlying problem. Our results suggest that academic policies and mentoring interventions targeting undergraduate advisors could contribute to reducing the gender disparity. Future research should evaluate the efficacy of educating faculty and students about the existence and impact of bias within academia, an approach that has reduced racial bias among students (43). Educational efforts might address research on factors that attenuate gender bias in real-world settings, such as increasing women's self-monitoring (44). Our results also point to the importance of establishing objective, transparent student evaluation and admissions criteria to guard against observers' tendency to unintentionally use different standards when assessing women relative to men (45, 46). Without such actions, faculty bias against female undergraduates may continue to undermine meritocratic advancement, to the detriment of research and education.

Conclusions

The dearth of women within academic science reflects a significant wasted opportunity to benefit from the capabilities of our best potential scientists, whether male or female. Although women have begun to enter some science fields in greater numbers (5), their mere increased presence is not evidence of the absence of bias. Rather, some women may persist in academic science despite the damaging effects of unintended gender bias on the part of faculty. Similarly, it is not yet possible to conclude that the preferences for other fields and lifestyle choices (9–11) that lead many women to leave academic science (even after obtaining advanced degrees) are not themselves influenced by experiences of bias, at least to some degree. To the extent that faculty gender bias impedes women's full participation in science, it may undercut not only academic meritocracy, but also the expansion of the scientific workforce needed for the next decade's advancement of national competitiveness (1).

Materials and Methods

Participants. We recruited faculty participants from Biology, Chemistry, and Physics departments at three public and three private large, geographically diverse research-intensive universities in the United States, strategically selected for their representative characteristics (see *SI Materials and Methods* for more information on department selection). The demographics of the 127 respondents corresponded to both the averages for the selected departments and faculty at all United States research-intensive institutions, meeting the criteria for generalizability even from nonrandom samples (see *SI Materials and Methods* for more information on recruitment strategy and participant characteristics). Indeed, we were particularly careful to obtain a sample representative of the underlying population, because many past studies have demonstrated that when this is the case, respondents and nonrespondents typically do not differ on demographic characteristics and responses to focal variables (47).

Additionally, in keeping with recommended practices, we conducted an a priori power analysis before beginning data collection to determine the optimal sample size needed to detect effects without biasing results toward obtaining significance (*SI Materials and Methods: Subjects and Recruitment Strategy*) (48). Thus, although our sample size may appear small to some readers, it is important to note that we obtained the necessary power and representativeness to generalize from our results while purposefully avoiding an unnecessarily large sample that could have biased our results toward a false-positive type I error (48).

Procedure. Participants were asked to provide feedback on the materials of an undergraduate science student who stated their intention to go on to graduate school, and who had recently applied for a science laboratory manager position. Of importance, participants believed they were evaluating a real student who would subsequently receive the faculty participants' ratings as feedback to help their career development (see *SI Materials and Methods* for more information, and Fig. S1 for the full text of the cover story). Thus, the faculty participants' ratings were associated with definite consequences.

Following established practices, the laboratory manager application was designed to reflect high but slightly ambiguous competence, allowing for variability in participant responses (20, 23). In addition, a promising but still-nascent applicant is precisely the type of student whose persistence in academic science is most likely to be affected by faculty support or discouragement (30–33), rendering faculty reactions to such a student of particular interest for the present purposes. The materials were developed in consultation with a panel of academic science researchers (who had extensive experience hiring and supervising student research assistants) to ensure that they would be perceived as realistic (*SI Materials and Methods*). Results of a funneled debriefing (49) indicated that this was successful; no participant reported suspicions that the target was not an actual student who would receive their evaluation.

Participants were randomly assigned to one of two student gender conditions: application materials were attributed to either a male student (John, $n = 63$), or a female student (Jennifer, $n = 64$), two names that have been pretested as equivalent in likability and recognizeability (50). Thus, each participant saw only one set of materials, from either the male or female applicant (see Fig. S2 for the full text of the laboratory manager application and *SI Method and Materials* for more information on all materials). Because all other information was held constant between conditions, any differences in participants' responses are attributable to the gender of the student. Using validated scales, participants rated student competence, their own likelihood of hiring the student, selected an annual starting salary for the student, indicated how much career mentoring they would provide to such a student, and completed the Modern Sexism Scale.

ACKNOWLEDGMENTS. We thank faculty members from six anonymous universities for their involvement as participants; and Jessamina Blum, John Crosnick, Jennifer Frederick, Jaime Napier, Jojanneke van der Toorn, Tiffany Tsang, Tessa West, James Young, and two anonymous reviewers for valuable input. This research was supported by a grant from the Howard Hughes Medical Institute Professors Program (to J.H.).

1. President's Council of Advisors on Science and Technology (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics.* Available at http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_feb.pdf. (Accessed February 13, 2012).
2. Handelsman J, et al. (2005) Careers in science. More women in science. *Science* 309:1190–1191.
3. United States National Academy of Sciences (2007) *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering* (National Academies, Washington, DC).
4. National Science Foundation (2009) *Women, Minorities, and Persons with Disabilities in Science and Engineering* (National Science Foundation, Arlington).
5. Bell N (2010) *Graduate Enrollment and Degrees: 1999 to 2009* (Council of Graduate Schools, Washington, DC).
6. Halpern DF, et al. (2007) The science of gender differences in science and mathematics. *Psychol Sci* 8(1):1–51.
7. Hyde JS, Linn MC (2006) Diversity. Gender similarities in mathematics and science. *Science* 314:599–600.
8. Spelke ES (2005) Sex differences in intrinsic aptitude for mathematics and science?: A critical review. *Am Psychol* 60:950–958.
9. Ceci SJ, Williams WM (2010) Gender differences in math-intensive fields. *Curr Dir Psychol Sci* 19:275–279.
10. Ceci SJ, Williams WM (2011) Understanding current causes of women's underrepresentation in science. *Proc Natl Acad Sci USA* 108:3157–3162.

11. Ceci SJ, et al. (2011) Do subtle cues about belongingness constrain women's career choices? *Psychol Inq* 22:255–258.
12. Berezow AB (2011) *Gender discrimination in science is a myth.* Available at http://www.nationalreview.com/articles/256816/gender-discrimination-science-myth-alex-b-berezow?pg=2. (Accessed June 6, 2012).
13. Dickey Zakaib G (2011) Science gender gap probed. *Nature* 470:153.
14. Devine PG (1989) Stereotypes and prejudice: Their automatic and controlled components. *J Pers Soc Psychol* 56(1):5–18.
15. Eagly AH, Mladinic A (1994) Are people prejudiced against women? Some answers from research on attitudes, gender stereotypes, and judgments of competence. *Eur Rev Soc Psychol* 5(1):1–35.
16. Spence JT, Hahn ED (1997) The attitudes toward women scale and attitude change in college students. *Psychol Women Q* 21(1):17–34.
17. Dovidio JF, Gaertner SL (2004) *Advances in Experimental Social Psychology*, ed Zanna MP (Elsevier, New York), pp 1–51.
18. Nosek BA, Banaji M, Greenwald AG (2002) Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dyn* 6(1):101–115.
19. Goldin C, Rouse C (2000) Orchestrating impartiality: The impact of "blind" auditions on female musicians. *Am Econ Rev* 90:715–741.
20. Foschi M (2000) Double standards for competence: Theory and research. *Annu Rev Sociol* 26(1):21–42.
21. Foschi M (1996) Double standards in the evaluation of men and women. *Soc Psychol Q* 59:237–254.
22. Goldberg P (1968) Are women prejudiced against women? *Transaction* 5(5):28–30.
23. Heilman ME, Wallen AS, Fuchs D, Tamkins MM (2004) Penalties for success: Reactions to women who succeed at male gender-typed tasks. *J Appl Psychol* 89:416–427.
24. Monin B, Miller DT (2001) Moral credentials and the expression of prejudice. *J Pers Soc Psychol* 81:33–43.
25. Uhlmann EL, Cohen GL (2007) "I think it, therefore it's true": Effects of self-perceived objectivity on hiring discrimination. *Organ Behav Hum Dec* 104:207–223.
26. Nosek BA, Banaji MR, Greenwald AG (2002) Math = male, me = female, therefore math not = me. *J Pers Soc Psychol* 83:44–59.
27. Dasgupta N (2011) Ingroup experts and peers as social vaccines who inoculate the self-concept: The stereotype inoculation model. *Psychol Inq* 22:231–246.
28. The Massachusetts Institute of Technology (1999) *A study on the status of women faculty in science at MIT.* Available at http://web.mit.edu/fnl/women/women.html#The%20Study (Accessed February 13, 2012).
29. Steele J, James JB, Barnett RC (2002) Learning in a man's world: Examining the perceptions of undergraduate women in male-dominated academic areas. *Psychol Women Q* 26(1):46–50.
30. Lent RW, et al. (2001) The role of contextual supports and barriers in the choice of math/science educational options: A test of social cognitive hypotheses. *J Couns Psychol* 48:474–483.
31. Seymour E, Hewitt NM (1996) *Talking about Leaving: Why Undergraduates Leave the Sciences* (Westview, Boulder).
32. Gasiewski JA, et al. (2012) From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Res Higher Educ* 53:229–261.
33. Byars-Winston A, Gutierrez B, Topp S, Carnes M (2011) Integrating theory and practice to increase scientific workforce diversity: A framework for career development in graduate research training. *CBE Life Sci Educ* 10:357–367.
34. Campbell DT, Fiske DW (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol Bull* 56:81–105.
35. Robins RW, Hendin HM, Trzesniewski KH (2001) Measuring global self esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Pers Soc Psychol Bull* 27(2):151–161.
36. Moss-Racusin CA, Rudman LA (2010) Disruptions in women's self-promotion: The backlash avoidance model. *Psychol Women Q* 34:186–202.
37. Rudman LA, Moss-Racusin CA, Glick P, Phelan JE (2012) *Advances in Experimental Social Psychology*, eds Devine P, Plant A (Elsevier, New York), pp 167–227.
38. Swim JK, Aikin KJ, Hall WS, Hunter BA (1995) Sexism and racism: Old-fashioned and modern prejudices. *J Pers Soc Psychol* 68:199–214.
39. Martell RF, Lane DM, Emrich C (1996) Male-female differences: A computer simulation. *Am Psychol* 51(2):157–158.
40. Steinpreis RE, Anders KA, Ritzke D (1999) The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles* 41:509–528.
41. Bandura A (1982) Self-efficacy mechanism in human agency. *Am Psychol* 37:122–147.
42. Scandura TA, Lankau MJ (1997) Relationships of gender, family responsibility and flexible work hours to organizational commitment and job satisfaction. *J Organ Behav* 18:377–391.
43. Rudman LA, Ashmore RD, Gary ML (2001) "Unlearning" automatic biases: The malleability of implicit prejudice and stereotypes. *J Pers Soc Psychol* 81:856–868.
44. O'Neill OA, O'Reilly CA (2011) Reducing the backlash effect: Self-monitoring and women's promotions. *J Occup Organ Psychol* 84:825–832.
45. Uhlmann EL, Cohen GL (2005) Constructed criteria: Redefining merit to justify discrimination. *Psychol Sci* 16:474–480.
46. Phelan JE, Moss-Racusin CA, Rudman LA (2008) Competent yet out in the cold: Shifting criteria for hiring reflects backlash toward agentic women. *Psychol Women Q* 32:406–413.
47. Holbrook AL, Krosnick JA, Pfent A (2007) *Advances in Telephone Survey Methodology*, eds Lepkowski JM, et al. (John Wiley & Sons, Hoboken), pp 499–528.
48. Simmons JP, Nelson LD, Simonsohn U (2011) False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* 22:1359–1366.
49. Bargh JA, Chartrand TL (2000) *Handbook of Research Methods in Social and Personality Psychology*, eds Reis HT, Judd CM (Cambridge Univ Press, New York), pp 253–285.
50. Brescoll VL, Uhlmann EL (2008) Can angry women get ahead? Status conferral, gender, and workplace emotion expression. *Psychol Sci* 19:268–275.
51. Cohen J (1998) *Statistical Power Analysis for the Behavioral Sciences* (Erlbaum, Hillsdale), 2nd Ed.

Adv Physiol Educ 30: 135–140, 2006;
doi:10.1152/advan.00087.2006.

Teaching and learning in the interactive classroom

D. U. Silverthorn

Department of Integrative Biology, University of Texas, Austin, Texas

Submitted 24 August 2006; accepted in final form 28 August 2006

Silverthorn, D. U. Teaching and learning in the interactive classroom. *Adv Physiol Educ* 30: 135–140, 2006; doi:10.1152/advan.00087.2006.—The Claude Bernard Distinguished Lectureship of the Teaching of Physiology Section is presented annually at the Experimental Biology meeting. The lectureship is named for Prof. Claude Bernard, the experimental physiologist who is credited with introducing the concept of homeostasis. The 2006 Claude Bernard Distinguished Lecture was given by Dr. Dee U. Silverthorn from the University of Texas at Austin, TX.

active learning; student-centered instruction; science education reform

IT IS A TREMENDOUS HONOR to give the Teaching Section's 2006 Claude Bernard Distinguished Lecture, and one of the nicest things about it is that it requires you to sit down and indulge in retrospection about your own teaching and that of those who came before. Consequently, I decided to start by learning more about Claude Bernard, the experimental physiologist after whom this lecture is named. Claude Bernard (1813–1878) is probably best known for his description of homeostasis–the constancy of the *milieu interieur*–but, as it turns out, Prof. Bernard could be considered an early proponent of the interactive classroom, the topic of my talk, because it was reported that he disliked lecturing and was much happier in the laboratory, where he could demonstrate physiological phenomena to his students and interact with them as he taught (3, 5).

For many centuries, the professor was the primary source of information, the font of knowledge. Books were nonexistent or scarce, as they still are today in developing countries of the world, and information was passed orally from teacher to pupil. The didactic lecture is an effective method for conveying information from one person to a larger number of students, but, as most of us have experienced, simply telling information to someone does not ensure that learning takes place. When I began teaching, the education paradigm had not changed much from the one that existed in the time of Claude Bernard. For example, an etching of a Smithsonian lecture hall circa 1856 (Fig. 1) shows the lecturer, or "sage on the stage," as the focus of the room, with students arranged in rows facing him. Even today, newly constructed classrooms often have the same configuration as those lecture halls built centuries ago.

In the last 30 years, however, technological advances have begun to change how students acquire facts. No longer do they need to depend on the teacher to tell them what they should know. Physiology textbooks have changed from page after page of printed text with a few simple black-and-white line drawings or graphs to glossy four-color publications with three-dimensional computer-aided illustrations that occupy more space than the text. I still remember my reaction upon seeing the first four-color physiology textbook to be published: "This is a comic book, not a serious textbook!" Now, textbooks come with a host of technology-driven ancillaries: interactive CDs, websites, animations, and simulated laboratory experiments . . . ways to convey information that were almost unimaginable 30 years ago.

Students have changed in the last 30 years as well. The generation of students we are now teaching grew up with computers. They have always known the internet, videos, and CDs, but they may have never seen a typewriter. They laugh when you tell them that computers used to be larger than cars, because the students in the class of 2007 have always had computers that fit into their backpacks (1). What this means is that we are teaching a generation whose view of information access and transfer is totally different from that of their older instructors. When students today want to know about something, they are far more likely to Google it or go to Wikipedia than they are to pull down a book from a bookshelf. Every year when I talk to my students about finding scientific information, at least two-thirds of my juniors and seniors have never gone into the stacks in one of the University of Texas libraries to look for a book. What is more depressing for older scientists whose publications predate electronic indexing in PubMed is that for many students, if they cannot find information online, it might as well not exist.

We, as teachers, must now recognize that our students no longer have to depend on us for the acquisition of information, which may be one reason some professors report low attendance in class. Why wake up for an 8 AM lecture if you can learn the material on your own? And that brings me to the fundamental question that we each need to answer: What is my role as a teacher? What can I do during my face-to-face time with students that they cannot do as effectively on their own?

We know that one thing that our students can do very well on their own is memorize facts. But, science education reform efforts in the last 10 years have been calling for teachers to move away from memorization of unrelated facts and instead emphasize better conceptual understanding of basic principles. However, progress in this arena has been very slow, particularly at the undergraduate level.

So, in the remainder of this discussion, I would like to examine three aspects of teaching in the 21st century that I believe support improved student learning. First, what happens in an interactive classroom, and how does it differ from the traditional lecture? Second, what happens to students when they come into an interactive classroom and are asked to change from passive note-taking mode to active participation? And, third, what happens to faculty when they either decide on their own or are told by the administration to change their teaching to a more interactive student-centered format?

Address for reprint requests and other correspondence: D. U. Silverthorn, Dept. of Integrative Biology, Univ. of Texas, 1 University Station, C0930, Austin, TX 78712 (e-mail: silverthorn@mail.utexas.edu).

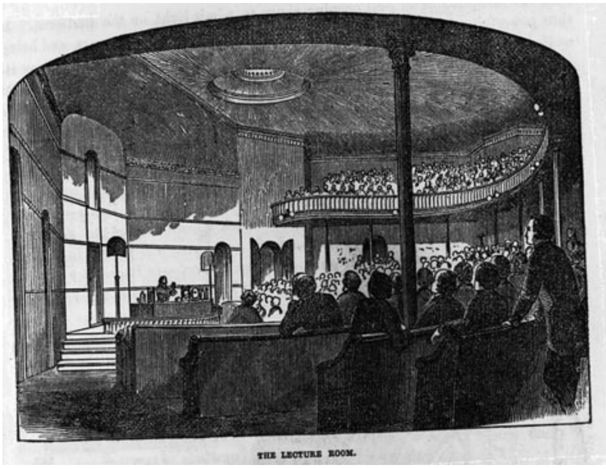


Fig. 1. A lecture hall in the Smithsonian Institute circa 1856 is similar to many classrooms in use today, with rows of students focused on the lecturer. [Reproduced with permission from the Smithsonian Institution Archives (record unit 95, box 31, folder 40, image no. MAH-43804k).]

The Interactive Classroom

As an example of an interactive classroom, I will describe what takes place in my large upper-division physiology classes at the University of Texas at Austin, TX (UT). My teaching strategy is not something that developed overnight, or even in one or two years. It has been a gradual process of trial and error. In many ways, my classroom has been my laboratory and my students are my lab rats. To quote Claude Bernard: "In experimentation . . . above all one must observe" (3). Each semester, I watch and listen to my students and my teaching assistants and try to decide which variable needs to be tweaked the next semester. Teaching is an iterative process, and I believe that when we stop trying to improve our teaching, it is time to retire.

Before I describe what takes place in my classroom, I want to be clear about two things. First, I am not advocating that we do away with lectures completely. There are some concepts in physiology (renal clearance comes to mind) that we all recognize as difficult for students to learn, and these concepts are probably best introduced in a lecture format in which the professor can ask and answer questions and monitor student understanding. Lectures are also essential for conveying the most recent scientific discoveries that may not yet be in textbooks, which are about a year out of date on the day when they first come out in print. But, I *am* advocating modifying the classic didactic lecture, in which the professor talks for 50 min or more, and perhaps asks a few questions that are answered by assertive students in the front rows, to a format that has students spending less time taking notes and more time testing their understanding of content.

I also need to say that I am not trying to advocate a single "best way" to teach. Teaching is highly individual and site specific, and what works for me in my classes may not work for you, particularly if you are not comfortable with it. For example, I have a colleague who specializes in classes where he takes on the persona of famous biologists through the centuries: Aristotle, Darwin, and Pasteur, for example. Aside from the fact that the vast majority of historically significant biologists before the 20th century were men, enacting classroom drama is just not my thing. We have to know ourselves and do what works best for us in our particular classroom.

In the traditional lecture class, students come to class, take notes on information that is given in the lecture, and then go home to study their notes, read the book (maybe), and work assigned homework problems. The difficulty with this strategy is that the teacher has no guarantee that the students learned anything during lecture or that they are learning at home. I recently gave a guest lecture on pulmonary gas exchange for a graduate class at a medical school. I was told that the students had heard lectures on pulmonary mechanics and oxygen transport, so I began class by asking them three simple questions on that material, using an electronic response system so that they could answer anonymously. Guess what? The majority of the class could not answer the questions correctly! After all, the test was still a week away. The teacher who had given the pulmonary physiology lectures was appalled. And how could I talk to them on pulmonary gas exchange if they did not remember the concepts underlying ventilation and hemoglobin binding of oxygen? That simple demonstration underscored an important lesson: just because we tell something to students does not mean they have learned it!

How can we tell what our students know and understand? The best way is to make them talk to us. From the time I started teaching I have always used a Socratic lecture format, where instead of simply conveying information for 50–90 min, I would pose questions for the class. But, I learned that even that simple form of interaction was threatening for many students in a large lecture setting. There was always a group of students, usually sitting at the front of the room, who would answer questions and talk to me as if we were chatting in my office, while the remainder of the class sat passively at the back and listened and took notes. So during the years I have been at UT, my class has evolved until now I barely lecture during a 90-min session. Most of my time with the students is spent having them talk and work on problems.

I have had many instructors tell me that they cannot give up lecture time because they have too much content to cover. But with my strategy, I have found that it is possible to convey significant amounts of content and make time for classroom activities. Creating a successful class like this requires five steps.

1. Develop clear objectives. In my classes, I decided that the development of basic skills is as important as learning content, so many of my class objectives are related to appropriate web searching, using indexes such as PubMed, reading and critiquing the scientific literature, and data analysis and presentation. Keeping these noncontent objectives in mind helps me design multipurpose classroom activities.

2. Identify essential content. This may be the single biggest stumbling block to changing the way science is taught in the United States. There has been such an exponential growth of what we know about biology since the 1980s that even researchers in a particular field are hard pressed to keep up with the literature. At the same time, in physiology and introductory biology, there has been a sense that we must teach it all. With each passing year, this becomes more difficult, yet many teachers are reluctant to cut back on the content they relay to their students. And they feel that the only way their students will learn is through a lecture format. I have actually heard teachers say, "They won't learn it if I don't tell it to them."

What happens when these students no longer have a teacher to tell them what to know?

One of the most valuable lessons I learned about identifying essential content came from working with a group of biomedical engineers as part of the National Science Foundation-sponsored Vanderbilt-Northwestern-Texas-Harvard (VaNTH) Engineering Research Center (www.vanth.org). Physiology is one of the core domains taught in biomedical engineering, but, as I talked to faculty in different programs, I realized that none of the engineering programs taught physiology as the traditional march through the physiological systems. Instead, I found that most programs concentrated on three to four systems, and they were not always the same ones. When I talked to the engineers about why they felt they could leave out certain physiological systems, they said they assumed that if their students had learned, for example, the basic concepts for fluid flow and pressure-flow relationships in the cardiovascular system, they could easily apply the same principles to air flow in the respiratory system. The key point here is that the students had to understand and remember the concepts, not simply memorize a bunch of equations and facts. Our task as teachers is to identify those essential concepts.

3. Decide what students can learn on their own. Can students learn basic facts on their own? Based on my teaching experience, I think that given well-written objectives and access to good resources, most students can teach themselves the basics. And I believe that by having the expectation that they will learn material on their own, we are fostering the skills and attitudes that they need to become self-directed life-long learners. The challenge of teaching this way is the student who comes to class with the attitude of "You're being paid to be the teacher . . . just tell me what I need to know."

To free up lecture time for working on problems, I make my students responsible for learning basic facts about a topic before they come to class. I decided that it was a waste of my time to stand up in lecture and say "The functions of the cardiovascular system are . . ." and wait while students wrote my list down. Some teachers speed up the notetaking process by giving the students copies of their Powerpoint slides, but then students may not come to class if everything they need to know is on the handout.

As a compromise, I created a student workbook that includes preclass reading assignments, information and figures for use in class, and lots and lots of problems. The preclass work tells the students which pages to read and includes basic content questions that are covered in the reading, such as "List the functions of the cardiovascular system" and "Trace a drop of blood from the left ventricle to the left atrium." If you have ever picked up a used textbook and seen the margin-to-margin yellow highlighting, you know how badly students need guidance on how to extract the key points from a paragraph. Some students answer the questions as they read, using the workbook to guide their note taking. Other students read the assignment and then test how well they understood what they read by trying to answer the questions. Student study strategies are as variable as teaching preferences, so I do not force them into any one method. But, I do expect them to have read and learned the basics before they come to class.

4. Use class time for practice and ungraded assessment. I usually start the class period with a brief overview of the topic for the day and perhaps a short quiz if I think they are not doing their preclass reading. Then, we move to asking and answering questions and doing small-group work. The student workbook contains the last three years' test questions, and I use those both for class problems and for additional practice. When we get to topics that I know are conceptually difficult, such as renal clearance, I may give a short lecture, but most of the time in class is spent working problems in groups.

The physical arrangement of the classroom is important for a successful interactive classroom. I teach in classrooms where the students can work comfortably with others around them and I can use my cordless microphone and walk between the rows. I usually roam the lecture hall, coming face to face with all the students . . . there is no place to hide, and everyone becomes accountable. I also use electronic response systems (Fig. 2) so that everyone answers the questions, not just the quickest or most vocal students. With these response systems, the students and I get instant feedback, and the teaching that takes place matches what the students need. Does a method like this work? I like to think so, and my evidence is that I have 95% attendance in an 8 AM class.

5. Make sure the graded assessment matches class activities. This may seem obvious, but I have observed several examples at my own institution where the assessment did not match the classroom activities, and student learning suffered as a result. In one instance, the professor used a traditional lecture setting to deliver very entertaining descriptions of classic experiments in biology and then wondered why the students could not design an experiment when presented a problem to solve on the test. In that instance, the teaching would have been more successful if the instructor during the lecture had given the class the question posed in the classic experiment and then allowed the students to brainstorm strategies for answering the question before describing how the experiment was actually conducted. In almost the reverse scenario, another teacher spent class time having students work on problems, but his tests focused on the memorization of trivial facts not covered in lecture. After two tests, many students stopped coming to class, and those that did attend talked about social matters during the problem-solving sessions because they had been trained that paying attention to what went on in class would not help them on the tests.

Student Reactions to the Interactive Classroom

Most students who attend an interactive class enjoy the challenge and working with their classmates. But, over the years, despite everything I have tried, there are always a few students who struggle through the semester and never make the transition from sitting passively in lecture to becoming an active participant. About 10 years ago that became the focus of my classroom research, which leads into my second topic: what happens to students when you ask them to participate in an interactive classroom?

This research started when I was working with a doctoral student in Science Education, Patti Thorn. Patti had a Master's in microbiology and was interested in active learning, so she decided to enroll in my Physiology course for one of her required science credits. After a few days in my class, she came to me extremely frustrated. Patti's prior education in science had been primarily through lectures and cookbook labs, and, despite the fact that she had studied educational theory in her

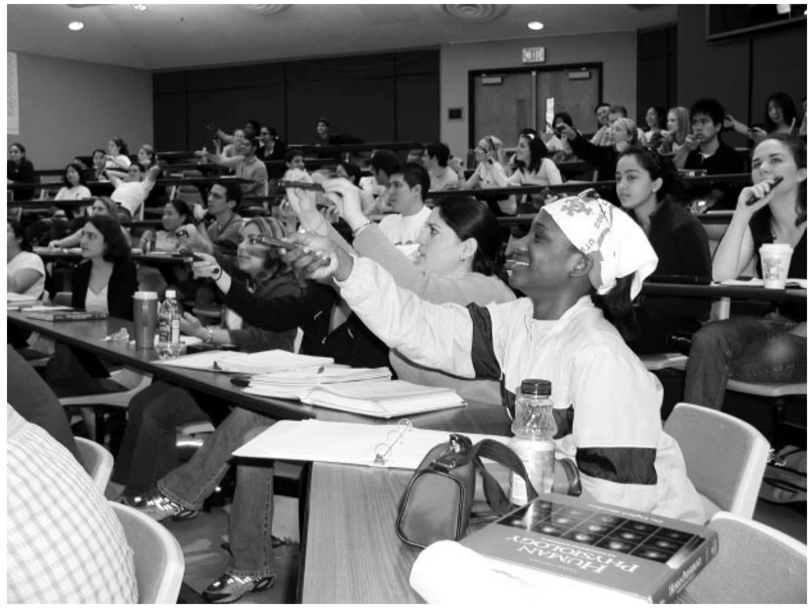


Fig. 2. Students in a large interactive classroom answer questions individually using an electronic response system.

doctoral program, she found that the reality of being in a class where she was expected to put learning theory into action was another matter. We talked about this conflict and agreed that Patti would chronicle her reactions to class in a journal. She also changed seats on a regular basis so that she could talk informally to the other students. Patti then shared the information she collected with me.

What I realized from analyzing my own and Patti's observations was that my students' behavior closely paralleled that reported by Donald Woods at McMaster University for students participating in a problem-based learning curriculum (7). When suddenly placed in a class that demands that they become independent learners, many students experience a fairly predictable series of reactions . . . reactions that are similar to Elisabeth Kubler-Ross's stages of coping with catastrophic news (4), although not necessarily in the same order.

The initial reaction is disbelief (denial). I observe this in my students on the first day of class when I tell them that this physiology class will not be the traditional lecture class they have come to expect. Typical comments include "Yeah, teachers always give you the idea that *their* class will be different and when it all shakes out, they are all the same." This stage persists for a few weeks until they realize that I am serious: I am not going to give them a lecture where they can write down what I want them to know, and they better not skip doing the workbook because if they do, they cannot follow what is going on during class. At that point, many students move into a second stage, which is shock or panic: "She is really serious . . . I can't believe this is happening."

The shock stage is quickly followed by what Woods called "strong emotion," which in my students manifests as a combination of anger and frustration. Typical student comments include "Why won't she just tell me what I need to know?" "She just needs to do her job and lecture!" and "Class time is worthless." What the students are really saying is "You've changed the rules!"

Most of my students are juniors and seniors with high grade-point averages, and by this stage in their college careers they have well-developed expectations of how a "good class" should be conducted. At UT, that often means they are expecting a well-organized, entertaining lecture where they can take copious notes that they then memorize to make a good grade on a multiple-choice exam. Many of these same students have never taken tests that do not give them a lot of content they recognize, and some have poorly developed thinking and reasoning skills. When they suddenly find themselves in a class where they can't make an A by the simple memorization of facts, and this may affect whether they get into medical school, they can become very hostile.

At this point, my main challenge is to overcome their resistance to change. A few students simply drop the class, but, for the ones who remain, it becomes very important to reassure them that the grading scheme in place will reward them at the end and not penalize them while they are trying to learn how to adapt to a new class style. In my class, this means a two-option grading scheme (Table 1), where the second option minimizes the weight of poor grades in the first part of the class. From here on, the students tend to follow one of two paths.

Most students accept the reality of the course structure and begin to adapt, particularly if they are flexible and can tolerate ambiguity. We have discovered that acceptance is the point

Table 1. *The two-option grading scheme for an interactive physiology course*

Option I	Option II
Three cumulative tests (75%)	Three cumulative tests (35%)
No final exam	Comprehensive final exam (40%)
Class work (10%)	Class work (10%)
Homework (10%)	Homework (10%)
Discussion attendance (5%)	Discussion attendance (5%)

For *option I,* students with a B average or better at the end of the semester may exempt from the final exam and take their *option I* grade. If the student takes the comprehensive final exam, then *option II* applies. *Option II* is required for everyone with less than a B average, including pass-fail students. Students who have an excused absence for a test must take the final exam and have the following grading scheme: comprehensive final exam (50%), class and discussion work (25%), and test average (25%).

where we have to be ready to help by encouraging students and letting them know that we have intentionally pushed them out of their "comfort zone." It is critical to give students alternative ways to approach the course, such as new study strategies. Most students, once their attempt to adapt meets with some success, experience a return of confidence. Often these students have to redefine what "success" means. Before this class, success was making an A on an exam. Now success is measured against progress ("I'm doing better than I was") and is related to mastery of the material.

The final stage for the students who adapt is a sense of self-empowerment. Often, the written course evaluations include comments such as "That was the hardest course I've had in college but I can't believe how much I've learned!" and "If I survived that . . . there's nothing I can't do!" These students become willing to tackle bigger goals with more confidence and such belief in themselves that it is hard to imagine they will be anything but successful. What is particularly interesting is that this last stage may not fully reach the students' consciousness until a year or two later, when they enter medical school and suddenly recognize how they have retained what they learned in their physiology class. I often learn about this stage by emails that begin with "You may not remember me . . ."

Unfortunately, there are usually a few students each semester who are unable to adapt and who continue to struggle with the demands of the class despite help. These students may become depressed, stop trying, or simply give up, saying "I can't learn this way." In some instances, continued failure to adapt to the new learning style causes the students to reexamine their career goals and decide that perhaps they should consider alternatives other than a career in medicine or biomedicine. If I can find the right intervention, some of them finally become successful. But, there are always a few students who never make the transition and fail.

Faculty and the Interactive Classroom

The challenge of coping with instructional change is not restricted to students. My experience working with a group of faculty members who were trying to incorporate more active learning in their classrooms (6) demonstrated that changing the way we teach is not simple. As with the students, there is a process and some critical barriers to overcome, and not everyone may be able to overcome them.

The reasons instructors have trouble changing how they teach are varied and complex. One simple reason is that many of us are products of the system that we are trying to change. We learned to teach with the "see it-do it" model, and, consequently, some faculty members have the attitude that "I learned this way; therefore, my students should be able to as well." Other factors that come into play are a lack of role models and a peer support system, lack of administrative support, and lack of appropriate teaching and assessment materials that deal with conceptual understanding and not simply memorization of facts. Finally, student resistance and anger, as discussed earlier, may impede the implementation of new teaching strategies. When teachers try something different in the classroom and students resist, the teacher may back down. Often, this is due to fear of what will happen to their student evaluations and contract renewals. I have been told by many instructors that they once tried active learning but the students hated it, so they went back to what was tried and true.

Successfully creating an interactive classroom requires a teacher who believes that students are capable of independent learning, given proper guidance and support. The interactive classroom becomes a place where learning focuses on concepts, principles, and application of knowledge rather than transfer of facts. In many ways, the classroom becomes where students learn what they do not know rather than what they do know.

So, here are six hints for success that emerged from my observations of students and faculty in interactive classrooms.

1. Define your goals and objectives. This step requires reflection on what we are teaching and to whom. We must be flexible enough to change our teaching to fit our student population and to tailor what we do to their needs. For example, I know that my prenursing students need more direction and hand-holding than my premed students do, but that my graduate students are not that different from my premeds. What is appropriate for one institution or population of students is not necessarily right for another.

2. Start small and don't change too many things at once. One of the biggest teaching disasters I have ever seen was a young postdoc who was teaching for the first time in our Nursing Physiology course. He had attended a faculty development seminar on student-centered teaching and enthusiastically decided to implement ALL the good ideas he heard about there. So he had his students working in teams, writing their own test questions, evaluating each other, and contracting with him for their grades. The one thing he did right was conduct a midsemester evaluation of his teaching, which told him that the students were all unhappy with the class. Unfortunately, there was no agreement as to which of his innovations was the worst, and he was left to salvage the semester as best he could.

3. Tell your students what you're doing and why, and KEEP TELLING THEM. This is one difference that we have noticed between faculty members who are successful in changing to an interactive classroom and those who continue to encounter student resistance. My story of why this is important comes from another colleague who came to one of our faculty development workshops and went home excited about student peer evaluation of written work. He started requiring his students to grade each other's laboratory reports and thought this technique had worked beautifully until he saw his teaching evaluations, which said, "The professor is lazy. He made us do his grading for him." His mistake was that he had failed to tell the students why he was having them grade each other's work.

Another colleague told students how the class structure would be different and why on the first day of class, but she did not repeat it. Remember the stages the students go through? When they hit the panic-anger stage, they have forgotten why you changed the rules on them. It is important to revisit your goals with students periodically so that they understand you are not teaching this way just to torture them.

Some years ago, I thought I had avoided this trap because at the beginning of the semester I talked to my students about how the class was going to be different from their usual UT science class. I showed them Bloom's taxonomy of educational objectives and told them how we would spend the semester concentrating on problem-solving and higher-level skills. I thought I was doing a great job of communicating my goals until I read one of my end of semester evaluations, which said, “I thought you said we weren’t going to have to memorize anything.” So, I changed what I say. Now, I tell students that they have to become a database of memorized information, stored and flexibly retrievable, so that they can find the information they need to solve problems they have never seen before.

4. Provide students with tools to help them change. How many times do students come in with an F on a test saying, “But I really knew the material!” It is important to teach students the difference between knowing and understanding and to show them that their study strategies should match the type of learning they desire. Many students develop study routines that made them successful in the “memorize and dump” classes, but when they find that their entrained study habits no longer work, they get frustrated. I use a variety of strategies to help these students make the transition to higher level learning. At the beginning of the semester, I have them take the visual-aural-read/write-kinesthetic learning preferences test with its study strategies (www.vark-learn.com), and I require them to make and use maps organizing large amounts of physiological information. Many of students initially resist these new ways of studying, but a lower grade on a test than they like is a powerful wakeup call and is often sufficient to initiate change.

5. Match the assessment to your teaching style, goals, and objectives. Assessment can be by group or individual. If your class time is spent problem solving but your tests demand memorization and regurgitation of petty details, students will decide that their time is better spent memorizing details. To show them that you value developing problem-solving skills, you must be consistent and give them classroom practice with the kinds of problems they will be asked to work on the test.

6. Have the right attitude. The final hint for success is for the teacher to approach classroom change with flexibility, patience, and a sense of humor. Usually nothing works the way you think it will, and sometimes it does not work at all and you need to rethink and try again. Finally, successful teachers constantly reflect on teaching and learning. This means thinking about each class . . . what worked, what didn’t. Where are the students having problems and what can I do to help them? This kind of reflection makes teaching a dynamic process, a creative endeavor.

We can all expect more challenges to change our teaching in the years to come. The iPod is almost ubiquitous now, and at some schools pod-casting lectures is becoming commonplace. I was talking to the director of a Medical Physiology course a couple of weeks ago, and his institution was considering taping the physiology lectures for students to view on their own time and then using the scheduled faculty contact time with the students for working problems and case studies. What other roles might teachers play in the future?

I would like to close with one more quote from Claude Bernard that I thought was particularly appropriate to this discussion: “A fact itself is nothing. It is valuable only for the ideas attached to it, or for the proof which it furnishes” (2). If, at the end of our course, the students appreciate the significance of this quote, then I think we have succeeded as teachers.

ACKNOWLEDGMENTS

I acknowledge Dr. Patti M. Thorn, whose insight contributed to my interpretation of what happens to students and faculty members who encounter or begin to implement interactive classroom techniques. I also thank all my graduate teaching assistants at the University of Texas whose input helped me decide which changes were working and which were not. I also acknowledge the support of Dr. Marilla Svinicki and the University of Texas Center for Teaching Excellence and Dr. Marsha Matyas and the American Physiological Society (APS) Education Office. Finally, thanks to my colleagues from the Teaching Section of the APS, the Human Anatomy and Physiology Society, and the International Union of Physiological Sciences.

REFERENCES

1. **Beloit College.** *Beloit College Mindset List. Beloit College’s Class of 2007 Mindset List* (online). http://www.beloit.edu/~pubaff/mindset/2007.htm [25 Mar 2006].
2. **Grande F and Visscher MB** (editors). *Claude Bernard and Experimental Medicine*. Cambridge, MA: Schenkman, 1967.
3. **Enersen OD.** *Who Named It? Whonamedit.com. Claude Bernard* (online). http://www.whonamedit.com/doctor.cfm/846.html [20 Mar 2006].
4. **Kubler-Ross E.** *On Death and Dying*. New York: MacMillan, 1969.
5. **Physiology Division, School of Biosciences, University of Westminster.** *Claude Bernard* (online). http://users.wmin.ac.uk/~mellerj/physiology/bernard.htm [20 Mar 2006].
6. **Silverthorn DU, Thorn PM, and Svinicki MD.** It’s difficult to change the way we teach: lessons from the Integrative Themes in Physiology Curriculum Module Project. *Adv Physiol Educ* 30:204–214, 2006.
7. **Woods DR.** *Problem-Based Learning: How to Gain the Most From PBL.* Waterdown, Ontario, Canada: Woods, 1994.

plored by combining models and further empirical data, but geology offers a stronger constraint because circumstances under which sulfate can be preserved in terrestrial sedimentary records are uncommon.

Although various aspects of Neoproterozoic glaciations are intensely disputed (*25*), our results confirm a profound difference from Phanerozoic ice ages. A near-global distribution of glaciated continents during the Marinoan phase ending ~635 million years ago is supported by evidence of low palaeomagnetic latitudes (*26*). The snowball Earth model (*27*) predicts a progressive accumulation of volcanic volatiles in the atmosphere that are not removed by weathering until the rapid demise of the ice age as the ice-albedo feedback reverses. If sulfate with large negative $\Delta^{17}O$ signals derived from oxidative weathering could only be generated in a large quantity after melting of the "snowball" and exposure of continents, then the diamictites above W2 had to be deposited during final glacial retreat, a hypothesis that should prompt a re-examination of their sedimentology. The alternative "slushball" model, in which parts of the ocean area are ice-free (*28*), would also permit accumulation of sulfate from prolonged oxidative weathering in certain continental "oases" where arid but cold conditions prevailed. This study provides an effective way to study the dynamics of sedimentation and atmospheric-hydrosphere-biosphere interactions during a global glaciation and highlights the need for further stratigraphically constrained $\Delta^{17}O_{SO4}$ data on continental carbonate precipitates to ground-truth flux-balance models.

References and Notes

1. H. Bao, D. Rumble, D. R. Lowe, *Geochim. Cosmochim. Acta* **71**, 4868 (2007).
2. $\delta^{18}O$ or $\delta^{17}O \equiv R^x_{sample}/R^x_{standard} - 1$ (where $R^x = {}^{18}O/{}^{16}O$ or ${}^{17}O/{}^{16}O$); the same δ notation applies to $\delta^{13}C$ or $\delta^{34}S$ in this paper.
3. Reference units for stable isotope compositions: VSMOW for sulfate $\delta^{18}O$, $\delta^{17}O$, and $\Delta^{17}O$; VPDB for carbonate $\delta^{13}C$ and $\delta^{18}O$; and Vienna Canyon Diablo Troilite for sulfate $\delta^{34}S$.
4. G. E. Claypool, W. T. Holser, I. R. Kaplan, H. Sakai, I. Zak, *Chem. Geol.* **28**, 199 (1980).
5. R. A. Socki, R. P. Harvey, D. L. Bish, E. Tonui, H. Bao, in *Lunar and Planetary Science Conference*, vol. XXXIX (NASA, Houston, TX, 2008), p. 1964.
6. H. Bao, *Chem. Geol.* **214**, 127 (2005).
7. H. Bao, M. H. Thiemens, D. B. Loope, X.-L. Yuan, *Geophys. Res. Lett.* **30**, 1843 (2003).
8. H. Bao, J. R. Lyons, C. M. Zhou, *Nature* **453**, 504 (2008).
9. G. P. Halverson, A. C. Maloof, P. F. Hoffman, *Basin Res.* **16**, 297 (2004).
10. G. P. Halverson, in *Neoproterozoic Geobiology and Paleobiology*, S. Xiao, A. J. Kaufman, Eds. (Springer, New York, 2006), pp. 231–271.
11. I. J. Fairchild, M. J. Hambrey, *Precambrian Res.* **73**, 217 (1995).
12. I. J. Fairchild, M. J. Hambrey, B. Spiro, T. H. Jefferson, *Geol. Mag.* **126**, 469 (1989).
13. Materials and methods are available as supporting material on *Science* Online.
14. G. W. Luther III, *Geochim. Cosmochim. Acta* **51**, 3193 (1987).
15. C. O. Moses, J. S. Herman, *Geochim. Cosmochim. Acta* **55**, 471 (1991).
16. N. Balci, W. C. Shanks, B. Mayer, K. W. Mandernack, *Geochim. Cosmochim. Acta* **71**, 3796 (2007).
17. M. R. Talbot, *Chem. Geol. Isol. Geol. Sect.* **80**, 261 (1990).
18. J. Farquhar, D. E. Canfield, A. Masterson, H. Bao, D. Johnston, *Geochim. Cosmochim. Acta* **72**, 2805 (2008).
19. N. F. Moreira, L. M. Walter, C. Vasconcelos, J. A. McKenzie, P. J. McCall, *Geology* **32**, 701 (2004).
20. I. J. Fairchild, M. J. Hambrey, *Precambrian Res.* **26**, 111 (1984).
21. B. J. Reedy, J. K. Beattie, R. T. Lowson, *Appl. Spectrosc.* **48**, 691 (1994).
22. A. Angert, S. Rachmilevitch, E. Barkan, B. Luz, *Global Biogeochem. Cycles* **17**, article 1030 (2003).
23. B. Luz, E. Barkan, *Geochim. Cosmochim. Acta* **69**, 1099 (2005).
24. R. T. Pierrehumbert, *J. Geophys. Res. Atmos.* **110**, article D01111 (2005).
25. I. J. Fairchild, M. J. Kennedy, *J. Geol. Soc. London* **164**, 895 (2007).
26. D. A. D. Evans, *Am. J. Sci.* **300**, 347 (2000).
27. P. F. Hoffman, D. P. Schrag, *Terra Nova* **14**, 129 (2002).
28. T. J. Crowley, W. T. Hyde, W. R. Peltier, *Geophys. Res. Lett.* **28**, 283 (2001).
29. H.B. and I.J.F. designed research and led the writing of the manuscript; H.B. performed CAS extraction and triple oxygen isotope measurements; I.J.F secured samples from field expeditions and conducted sedimentological, petrographic, mineralogical and elemental studies; P.M.W. conducted preliminary CAS extraction and performed $\delta^{34}S_{CAS}$ analysis; and C.S. carried out $\delta^{13}C$ and $\delta^{18}O$ analysis of host carbonates. We thank G. Halverson for discussion and Y. Peng for analytical assistance. Financial and facility supports were provided by Louisiana State University, NSF, and Chinese Academy of Science (H.B.), Natural Environment Research Council (NERC) standard grant (GR3/C511805/1) and NERC inductively coupled plasma mass spectrometry facilities (I.J.F.), and Austrian Science Funds (C.S.). The authors declare no competing financial interests.

Supporting Online Material

www.sciencemag.org/cgi/content/full/323/5910/119/DC1
Materials and Methods
SOM Text
Figs. S1 and S2
Tables S1 and S2
References

3 September 2008; accepted 25 November 2008
10.1126/science.1165373

Why Peer Discussion Improves Student Performance on In-Class Concept Questions

M. K. Smith,[1]* W. B. Wood,[1] W. K. Adams,[2] C. Wieman,[2,3] J. K. Knight,[1] N. Guild,[1] T. T. Su[1]

When students answer an in-class conceptual question individually using clickers, discuss it with their neighbors, and then revote on the same question, the percentage of correct answers typically increases. This outcome could result from gains in understanding during discussion, or simply from peer influence of knowledgeable students on their neighbors. To distinguish between these alternatives in an undergraduate genetics course, we followed the above exercise with a second, similar (isomorphic) question on the same concept that students answered individually. Our results indicate that peer discussion enhances understanding, even when none of the students in a discussion group originally knows the correct answer.

In undergraduate science courses, conceptual questions that students answer using personal response systems or "clickers" are promoted as a means to increase student learning [e.g. (*1*, *2*)], often through peer instruction (PI) (*3*). Instructors using this approach break up their lectures with multiple-choice questions to test understanding of the concepts being presented. When PI is used, students are first asked to answer a question individually, and then a histogram of their responses may be displayed to the class. If there is substantial disagreement among responses, students are invited to discuss questions briefly with their neighbors and then revote before the correct answer is revealed. The instructor then displays the new histogram and explains the reasoning behind the correct answer. Most instructors report that the percentage of correct answers, as well as students' confidence in their answers, almost always increases after peer discussion (*2*–*4*).

It is generally assumed that active engagement of students during discussion with peers, some of whom know the correct answer, leads to increased conceptual understanding, resulting in improved performance after PI. However, there is an alternative explanation: that students do not in fact learn from the discussion, but simply choose the answer most strongly supported by neighbors they perceive to be knowledgeable. We sought to distinguish between these alternatives, using an additional, similar clicker question that students answered individually to test for gains in understanding. Our results indicate that peer discussion enhances understanding, even when none of the students in a discussion group originally knows the correct answer.

In an undergraduate introductory genetics course for biology majors at the University of Colorado–Boulder (additional demographic in-

[1]Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309, USA. [2]Department of Physics, University of Colorado, Boulder, CO 80309, USA. [3]Department of Physics, University of British Columbia, Vancouver, BC V6T 1Z3, Canada.

*To whom correspondence should be addressed. E-mail: michelle.k.smith@colorado.edu

formation in table S1), we asked an average of five clicker questions per 50-min class throughout the semester and encouraged students to discuss questions with their neighbors. Students were given participation points for answering clicker questions, regardless of whether their answers were correct. Exam questions were similar to the clicker questions, so that students had an incentive to take clicker questions seriously.

Sixteen times during the semester we assessed how much students learned from peer discussion by using a paired set of similar (isomorphic) clicker questions. Isomorphic questions have different "cover stories," but require application of the same principles or concepts for solution (*5*, *6*). Sample isomorphic question pairs are shown in fig. S1. In class, students were first asked to answer one question of the pair individually (Q1). Then they were invited to discuss the question with their neighbors and revote on the same question ($Q1_{ad}$ for "Q1 after discussion"). Finally, students were asked to answer the second isomorphic question, again individually (Q2). Neither the answers to the two questions ($Q1/Q1_{ad}$ and Q2) nor the histograms of student answers were revealed until after the voting on Q2, so that there was minimal instructor or whole-course peer influence on the Q2 responses. The isomorphic questions were randomly assigned as $Q1/Q1_{ad}$ or Q2 after both questions were written. Data analysis was limited to students who answered all three questions of an isomorphic pair with a total of 350 students participating in the study (*7*) (see supporting online text).

Two results indicate that most students learned from the discussion of Q1. First, using data pooled from individual mean scores on Q1, $Q1_{ad}$, and Q2 for all 16 question pairs, the average percentage correct for Q2 was significantly higher than for Q1 and $Q1_{ad}$ (Fig. 1A and Table 1). Second, of the students who answered Q1 incorrectly and $Q1_{ad}$ correctly, 77% answered Q2 correctly (Fig. 2). This result suggests that most students who initially did not understand a concept were able to apply information they learned during the group discussion and correctly answer an isomorphic question. In contrast, almost all students who answered Q1 correctly, presumably because they understood the concept initially, did not change their votes on $Q1_{ad}$ and went on to answer Q2 correctly (Fig. 2).

In addition, students who answered both Q1 and $Q1_{ad}$ incorrectly still appeared to learn from discussions with peers and answering a second question on the same topic. Of these students, 44% answered Q2 correctly, significantly better than expected from random guessing (Fig. 2; on average, the questions in our 16 isomorphic pairs had four answer choices each). This result was unexpected because when students answered Q2, they had not been told the correct answer to $Q1/Q1_{ad}$, had not seen histograms of student responses, and had not discussed Q2 with their peers. We speculate that when this group of students discussed Q1, they were making sense of the information, but were unable to apply their new knowledge until presented with a fresh question on the same concept (Q2). There may also be a learning benefit to considering successive clicker questions on the same topic (*8*).

Although the difficulty of the question pairs varied, as judged by the percentage of correct answers on Q1 (see supporting online text), students performed significantly better on $Q1_{ad}$ and Q2 compared to Q1 for each difficulty level (Fig. 1B and Table 1). On the most difficult questions there was another significant increase between $Q1_{ad}$ and Q2, suggesting that there was an additional delayed benefit to the group discussions.

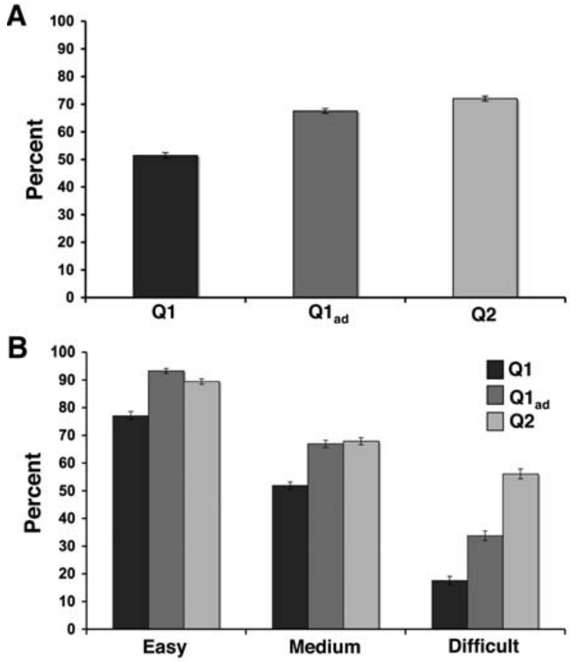


Fig. 1. The percentage of students who can correctly answer a question as individuals increases after peer discussion of a similar (isomorphic) question. Q1: One question of an isomorphic pair was voted on individually; $Q1_{ad}$: the same question was voted on again after peer discussion; Q2: the second isomorphic question was voted on individually. (**A**) Results for all 16 question pairs were averaged for each individual (n = 350 students), and the class averages of these scores are shown. (**B**) The 16 paired questions were grouped according to difficulty based on the percentage of correct answers for Q1 (five easy questions, seven medium questions, and four difficult questions), and performance results were again averaged for each individual (n = 343 students for easy, 344 for medium, and 337 for difficult) before computing the averages shown. Error bars show the SEM.

Table 1. Mean differences between Q1, $Q1_{ad}$, and Q2. The SEM is in parentheses.

Question category	$Q1_{ad} - Q1$* (%)	$Q2 - Q1$* (%)	$Q2 - Q1_{ad}$* (%)
All questions	16(1)	21(1)	5(1)
Easy questions	16(1)	12(2)	−4(1)†
Medium questions	15(1)	16(2)	1(1)†
Difficult questions	16(2)	38(2)	22(2)

*Mean values are the averages of the differences between $Q1_{ad}$-Q1, Q2-Q1, and Q2-$Q1_{ad}$ for each student. †No significant improvement between these questions.

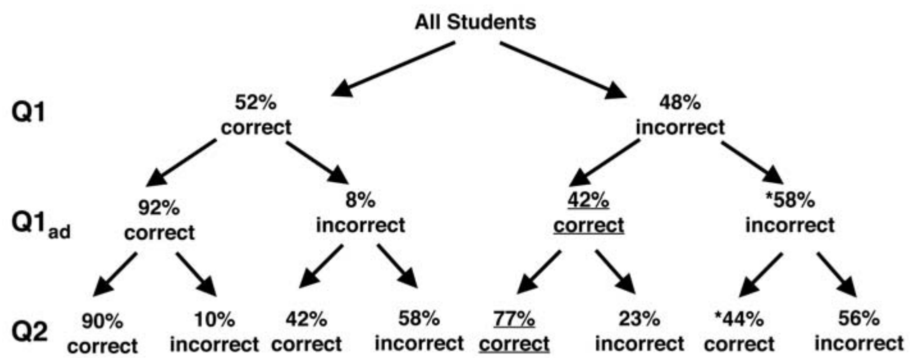


Fig. 2. Breakdown of student responses for the pool of 16 Q1, $Q1_{ad}$, and Q2 questions. Percentages of the category are connected by arrows from the preceding line. Underlined entries represent students who initially did not answer Q1 correctly but did so after group discussion; entries with an asterisk represent students who did not answer either Q1 or $Q1_{ad}$ correctly, but nevertheless were able to correctly answer the isomorphic question Q2. Of the 32 questions in our 16 question pairs, 7 had 5 answer choices, 5 had 4 choices, 3 had 3 choices, and 1 had 2 choices.

Our results suggest that peer discussion can be effective for understanding difficult concepts even when no one in the group initially knows the correct answer. In a postsemester survey (n = 98 responding), students reported an average of three participants in their peer discussion groups. If students who knew the answer to Q1 were randomly distributed throughout the classroom, then on the difficult questions (Fig. 1B), more than half of the 84 groups would have included no one who knew the correct answer to Q1 (naïve groups). Statistical analysis (see supporting online text) shows that some students who answered Q2 correctly must have come from naïve groups.

Student opinion supported the view that having someone in the group who knows the correct answer is unnecessary. On an end-of-year survey (n = 328 responding), 47% of students disagreed with the statement: "When I discuss clicker questions with my neighbors, having someone in the group who knows the correct answer is necessary in order to make the discussion productive." Representative comments from these students included the following: "Often when talking through the questions, the group can figure out the questions without originally knowing the answer, and the answer almost sticks better that way because we talked through it instead of just hearing the answer." "Discussion is productive when people do not know the answers because you explore all the options and eliminate the ones you know can't be correct."

This study supports the substantial value of student peer discussion as an effective means of active learning in a lecture class. Our findings are consistent with earlier demonstrations of social learning, including the value of discussion with peers (*9*–*13*). The significant increases in performance between Q1 and $Q1_{ad}$ confirm results from earlier classroom studies (*2*–*4*). In addition, we have presented new evidence showing that these increases result primarily from student gains in conceptual understanding rather than simply from peer influence.

Previous explanations for the value of PI have maintained the "transmissionist" view (*14*) that during discussion, students who know the right answer are explaining the correct reasoning to their less knowledgeable peers, who consequently improve their performance on the revote (*3*, *4*). Our finding that even students in naïve groups improve their performance after discussion suggests a more constructivist explanation: that these students are arriving at conceptual understanding on their own, through the process of group discussion and debate.

Some instructors who use clicker questions skip peer discussion entirely, believing that instructor explanation of the correct reasoning will be more clear and accurate than an explanation by peers, and will therefore lead to more student learning. Although our current work does not directly compare the benefits of instructor versus peer explanation, research in physics has shown that instructor explanations often fail to produce gains in conceptual understanding (*15*). We have shown that peer discussion can effectively promote such understanding. Furthermore, justifying an explanation to a fellow student and skeptically examining the explanation of a peer provide valuable opportunities for students to develop the communicative and metacognitive skills that are crucial components of disciplinary expertise.

References and Notes

1. D. Duncan, *Clickers in the Astronomy Classroom* (Pearson Education, San Francisco, 2006).
2. J. K. Knight, W. B. Wood, *Cell Biol. Educ.* **4**, 298 (2005).
3. E. Mazur, *Peer Instruction: A User's Manual* (Prentice Hall, Saddle River, NJ, 1997).
4. C. H. Crouch, E. Mazur, *Am. J. Phys.* **69**, 970 (2001).
5. K. Kotovsky, J. Hayes, H. Simon, *Cognit. Psychol.* **17**, 248 (1985).
6. H. Simon, J. Hayes, *Cognit. Psychol.* **8**, 165 (1976).
7. Internal Review Board, University of Colorado, Boulder, Approval to evaluate clicker responses (exempt status, Protocol No. 0108.9).
8. N. W. Reay, P. F. Li, L. Bao, *Am. J. Phys.* **76**, 171 (2008).
9. P. Cohen, J. Kulik, C.-L. Kulik, *Am. Educ. Res. J.* **19**, 237 (1982).
10. M. T. H. Chi, N. De Leeuw, M. H. Chiu, C. LaVancher, *Cogn. Sci.* **18**, 439 (1994).
11. N. Webb, L. Cullian, *Am. Educ. Res. J.* **20**, 411 (1983).
12. A. S. Palinscar, A. L. Brown, *Cogn. Instr.* **1**, 117 (1984).
13. E. B. Coleman, I. D. Rivkin, A. L. Brown, *J. Learn. Sci.* **6**, 347 (1997).
14. We use this term to describe the view that learning during instruction occurs by transmission of information from a teacher to a learner.
15. D. Hestenes, *Phys. Teach.* **30**, 141 (1992).
16. M.K.S., W.K.A. and J.K.K. were supported by the University of Colorado Science Education Initiative.

Supporting Online Material
www.sciencemag.org/cgi/content/full/323/5910/122/DC1
SOM Text
Fig. S1
Tables S1 to S3

15 September 2008; accepted 10 November 2008
10.1126/science.1165919

Regulation of Neuronal Survival Factor MEF2D by Chaperone-Mediated Autophagy

Qian Yang,[1] Hua She,[1] Marla Gearing,[2] Emanuela Colla,[3] Michael Lee,[3] John J. Shacka,[4] Zixu Mao[1,2]*

Chaperone-mediated autophagy controls the degradation of selective cytosolic proteins and may protect neurons against degeneration. In a neuronal cell line, we found that chaperone-mediated autophagy regulated the activity of myocyte enhancer factor 2D (MEF2D), a transcription factor required for neuronal survival. MEF2D was observed to continuously shuttle to the cytoplasm, interact with the chaperone Hsc70, and undergo degradation. Inhibition of chaperone-mediated autophagy caused accumulation of inactive MEF2D in the cytoplasm. MEF2D levels were increased in the brains of α-synuclein transgenic mice and patients with Parkinson's disease. Wild-type α-synuclein and a Parkinson's disease–associated mutant disrupted the MEF2D-Hsc70 binding and led to neuronal death. Thus, chaperone-mediated autophagy modulates the neuronal survival machinery, and dysregulation of this pathway is associated with Parkinson's disease.

In neurodegenerative diseases, certain populations of adult neurons are gradually lost because of toxic stress. The four myocyte enhancer factor 2 (MEF2) transcription factors, MEF2A to MEF2D, have been shown to play an important role in the survival of several types of neurons, and a genetic polymorphism of the MEF2A gene has been linked to the risk of late onset of Alzheimer's disease (*1*–*3*). In cellular models, inhibition of MEF2s contributes to neuronal death. Enhancing MEF2 activity protects neurons from death in vitro and in the substantia nigra pars compacta in a mouse model of Parkinson's disease (PD) (*4*). Neurotoxic insults cause MEF2 degradation in part by a caspase-dependent mechanism (*5*), but how MEF2 is regulated under basal conditions without overt toxicity is unknown. Autophagy refers to the degradation of intracellular components by lysosomes. Relative to macro- and microautophagy, chaperone-mediated autophagy (CMA) selectively degrades cytosolic proteins (*6*). This process involves binding of heat shock protein Hsc70 to substrate proteins via a KFERQ-like motif and their subsequent targeting to lysosomes via the lysosomal membrane receptor Lamp2a. Dysregulation of autophagy plays a role in neurodegeneration (*7*–*9*). However, the direct mechanism by which CMA modulates neuronal survival or death is unclear.

[1]Department of Pharmacology, Emory University School of Medicine, Atlanta, GA 30322, USA. [2]Department of Neurology, Emory University School of Medicine, Atlanta, GA 30322, USA. [3]Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. [4]Department of Pathology, Division of Neuropathology, University of Alabama at Birmingham, Birmingham, AL 35294, USA.

*To whom correspondence should be addressed. E-mail: zmao@pharm.emory.edu

Chapter 1

Backward Design

Design, v.,—To have purposes and intentions; to plan and execute
—*Oxford English Dictionary*

The complexity of design work is often underestimated. Many people believe they know a good deal about design. What they do not realize is how much more they need to know to do design well, with distinction, refinement, and grace.
—John McClean, "20 Considerations That Help a Project Run Smoothly," 2003

Teachers are designers. An essential act of our profession is the crafting of curriculum and learning experiences to meet specified purposes. We are also designers of assessments to diagnose student needs to guide our teaching and to enable us, our students, and others (parents and administrators) to determine whether we have achieved our goals.

Like people in other design professions, such as architecture, engineering, or graphic arts, designers in education must be mindful of their audiences. Professionals in these fields are strongly client-centered. The effectiveness of their designs corresponds to whether they have accomplished explicit goals for specific end-users. Clearly, students are our primary clients, given that the effectiveness of curriculum, assessment, and instructional designs is ultimately determined by their achievement of desired learnings. We can think of our designs, then, as software. Our courseware is designed to make learning more effective, just as computer software is intended to make its users more productive.

As in all the design professions, standards inform and shape our work. The software developer works to maximize user-friendliness and to reduce bugs that impede results. The architect is guided by building codes, customer budget, and neighborhood aesthetics. The teacher as designer is similarly constrained. We are not free to teach any topic we choose by any means. Rather, we are guided by national, state, district, or institutional standards that specify what students should know and be able to do. These standards provide a useful framework to help us identify teaching and learning priorities and guide our design of curriculum and assessments. In addition to external standards, we must also factor in the needs of our many and varied students when designing learning experiences. For example, diverse student interests, developmental levels, large classes, and previous achievements must always shape our thinking about the learning activities, assignments, and assessments.

Yet, as the old adage reminds us, in the best designs form follows function. In other words, all the methods and materials we use are shaped by a clear conception of the vision of desired results. That means that we must be able to state with clarity what the student should understand and be able to do as a result of any plan and irrespective of any constraints we face.

You probably know the saying, "If you don't know exactly where you are headed, then any road will get you there." Alas, the point is a serious one in education. We are quick to say what things *we* like to teach, what activities *we* will do, and what kinds of resources *we* will use; but without clarifying the desired results of our teaching, how will we ever know whether our designs are appropriate or arbitrary? How will we distinguish merely interesting learning from *effective* learning? More pointedly, how will we ever meet content standards or arrive at hard-won student understandings unless we think through what those goals imply for the learner's activities and achievements?

Good design, then, is not so much about gaining a few new technical skills as it is about learning to be more thoughtful and specific about our purposes and what they imply.

Why "backward" is best

How do these general design considerations apply to curriculum planning? Deliberate and focused instructional design requires us as teachers and curriculum writers to make an important shift in our thinking about the nature of our job. The shift involves thinking a great deal, first, about the specific learnings sought, and the evidence of such learnings, before thinking about what we, as the teacher, will do or provide in teaching and learning activities. Though considerations about what to teach and how to teach it may dominate our thinking as a matter of habit, the challenge is to focus first on the desired learnings from which appropriate teaching will logically follow.

Our lessons, units, and courses should be logically inferred from the results sought, not derived from the methods, books, and activities with which we are most comfortable. Curriculum should lay out the most effective ways of achieving specific results. It is analogous to travel planning. Our frameworks should provide a set of itineraries deliberately designed to meet cultural goals rather than a purposeless tour of all the major sites in a foreign country. In short, the best designs derive backward from the learnings sought.

The appropriateness of this approach becomes clearer when we consider the educational purpose that is the focus of this book: understanding. We cannot say *how* to teach for understanding or *which* material and activities to use until we are quite clear about which specific understandings we are after and what such understandings look like in practice. We can best decide, as guides, what "sites" to have our student "tourists" visit and what specific "culture" they should experience in their brief time there only if we are clear about the particular understandings about the culture we want them to take home. Only by having specified the desired results can we focus on the content, methods, and activities most likely to achieve those results.

But many teachers begin with and remain focused on textbooks, favored lessons, and time-honored activities—the inputs—rather than deriving those means from what is implied in the desired results—the output. To put it in an odd way, too many teachers focus on the *teaching* and not the *learning*. They spend most of their time thinking, first, about what they will do, what materials they will use, and what they will ask students to do rather than first considering what the learner will need in order to accomplish the learning goals.

Consider a typical episode of what might be called *content*-focused design instead of *results*-focused design. The teacher might base a lesson on a particular topic (e.g., racial prejudice), select a resource (e.g., *To Kill a Mockingbird*), choose specific instructional methods based on the resource and topic (e.g., Socratic seminar to discuss the book and cooperative groups to analyze stereotypical images in films and on television), and hope thereby to cause learning (and meet a few English/language arts standards). Finally, the teacher might think up a few essay questions and quizzes for assessing student understanding of the book.

This approach is so common that we may well be tempted to reply, What could be wrong with such an approach? The short answer lies in the basic questions of purpose: Why are we asking students to read this particular novel—in other words, what *learnings* will we seek from their having read it? Do the students grasp why and how the purpose should influence their studying? What should students be expected to understand and do upon reading the book, related to our goals beyond the book? Unless we begin our design work with a clear insight into larger purposes—whereby the book is properly thought of as a means to an educational end, not an end unto itself—it is unlikely that all students will *understand* the book (and their performance obligations). Without being self-conscious of the specific understandings about prejudice we seek, and how reading and discussing the book will help develop such insights, the goal is far too vague: The approach is more "by hope" than "by design." Such an approach ends up unwittingly being one that could be described like this: Throw some content and activities against the wall and hope some of it sticks.

Design Tip

Consider these questions that arise in the minds of all readers, the answers to which will frame the priorities of coached learning: How should I read the book? What am I looking for? What will we discuss? How should I prepare for those discussions? How do I know if my reading and discussions are effective? Toward what performance goals do this reading and these discussions head, so that I might focus and prioritize my studies and note taking? What big ideas, linked to other readings, are in play here? These are the students' proper questions about the learning, not the teaching, and any good educational design answers them from the start and throughout a course of study with the use of tools and strategies such as graphic organizers and written guidelines.

Answering the "why?" and "so what?" questions that older students always ask (or want to), and doing so in concrete terms as the focus of curriculum planning, is thus the essence of understanding by design. What is difficult for many teachers to see (but easier for students to feel!) is that, without such explicit and transparent priorities, many students find day-to-day work confusing and frustrating.

The twin sins of traditional design

More generally, weak educational design involves two kinds of purposelessness, visible throughout the educational world from kindergarten through graduate school, as noted in the Introduction. We call these the "twin sins" of traditional design. The error of activity-oriented design might be called "hands-on without being minds-on"—engaging experiences that lead only accidentally, if at all, to insight or achievement. The activities, though fun and interesting, do not lead anywhere intellectually. As typified by the apples vignette in the Introduction, such activity-oriented curricula lack an explicit focus on important ideas and appropriate evidence of learning, especially in the minds of the learners. They think their job is merely to engage; they are led to think the learning *is* the activity instead of seeing that the learning comes from being asked to consider the *meaning* of the activity.

A second form of aimlessness goes by the name of "coverage," an approach in which students march through a textbook, page by page (or teachers through lecture notes) in a valiant attempt to traverse all the factual material within a prescribed time (as in the world history vignette in the Introduction). Coverage is thus like a whirlwind tour of Europe, perfectly summarized by the old movie title *If It's Tuesday, This Must Be Belgium,* which properly suggests that no overarching goals inform the tour.

As a broad generalization, the activity focus is more typical at the elementary and lower middle school levels, whereas coverage is a prevalent secondary school and college problem. Yet, though the apples and world history classrooms look quite different with lots of physical activity and chatter in the former versus lecturing and quiet note taking in the latter, the design result is the same in both cases: No guiding intellectual purpose or clear priorities frame the learning experience. In neither case can students see and answer such questions as these: What's the point? What's the big idea here? What does this help us understand or be able to do? To what does this relate? Why should we learn this? Hence, the students try to engage and follow as best they can, hoping that meaning will emerge.

MISCONCEPTION ALERT!

Coverage is not the same as *purposeful survey.* Providing students with an overview of a discipline or a field of study is not inherently wrong. The question has to do with the transparency of purpose. *Coverage* is a negative term (whereas *introduction* or *survey* is not) because when content is "covered" the student is led through unending facts, ideas, and readings with little or no sense of the overarching ideas, issues, and learning goals that might inform study. (See Chapter 10 for more on coverage versus uncoverage.)

Students will be unable to give satisfactory responses when the design does not provide them with clear purposes and explicit performance goals highlighted throughout their work. Similarly, teachers with an activity or coverage orientation are less likely to have acceptable answers to the key design questions: What should students understand as a result of the activities or the content covered? What should the experiences or lectures equip them to do? How, then, should the activities or class discussions be shaped and processed to achieve the desired results? What would be evidence that learners are en route to the desired abilities and insights? How, then, should all activities and resources be chosen and used to ensure that the learning goals are met and the most appropriate evidence produced? How, in other words, will students be helped to see *by design* the purpose of the activity or resource and its helpfulness in meeting specific performance goals?

We are advocating the reverse of common practice, then. We ask designers to start with a much more careful statement of the desired results—the priority *learnings*—and to derive the curriculum from the performances called for or implied in the goals. Then, contrary to much common practice, we ask designers to consider the following questions after framing the goals: What would count as evidence of such achievement? What does it look like to meet these goals? What, then, are the implied *performances* that should make up the assessment, toward which all teaching and learning should point? Only after answering these questions can we logically derive the appropriate teaching and learning experiences so that students might perform successfully to meet the standard. The shift, therefore, is away from starting with such questions as "What book will we read?" or "What activities will we do?" or "What will we discuss?" to "What should they walk out the door able to understand, regardless of what activities or texts we use?" and "What is evidence of such ability?" and, therefore, "What texts, activities, and methods will best enable such a result?" In teaching students for understanding, we must grasp the key idea that *we are coaches of their ability to play the "game" of performing with understanding, not tellers of our understanding to them on the sidelines.*

Design Tip

To test the merits of our claims about purposelessness, we encourage you to sidle up to a student in the middle of any class and ask the following questions:

What are you doing?

Why are you being asked to do it?

What will it help you do?

How does it fit with what you have previously done?

How will you show that you have learned it?

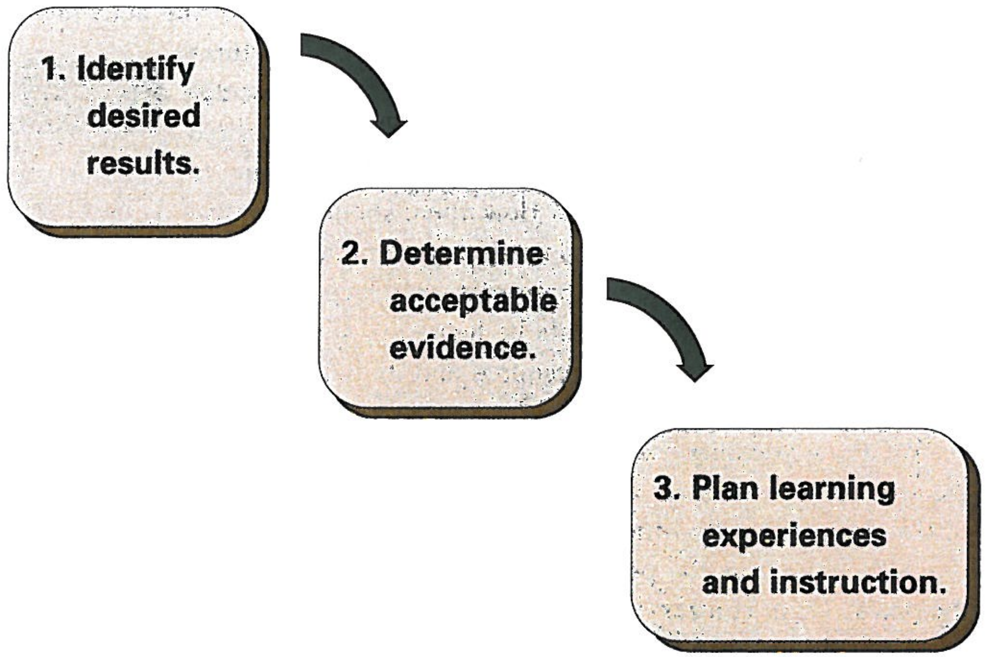
The three stages of backward design

We call this three-stage approach to planning "backward design." Figure 1.1 depicts the three stages in the simplest terms.

Stage 1: Identify desired results

What should students know, understand, and be able to do? What content is worthy of understanding? What *enduring* understandings are desired?

Figure 1.1
UbD: Stages of Backward Design



In Stage 1 we consider our goals, examine established content standards (national, state, district), and review curriculum expectations. Because typically we have more content than we can reasonably address within the available time, we must make choices. This first stage in the design process calls for clarity about priorities.

Stage 2: Determine acceptable evidence

How will we know if students have achieved the desired results? What will we accept as evidence of student understanding and proficiency? The backward design orientation suggests that we think about a unit or course in terms of the collected assessment evidence needed to document and validate that the desired learning has been achieved, not simply as content to be covered or as a series of learning activities. This approach encourages teachers and curriculum planners to first "think like an assessor" before designing specific units and lessons, and thus to consider up front how they will determine if students have attained the desired understandings.

Stage 3: Plan learning experiences and instruction

With clearly identified results and appropriate evidence of understanding in mind, it is now the time to fully think through the most appropriate instructional activities. Several key questions must be considered at this stage of backward design: What enabling knowledge (facts, concepts, principles) and skills (processes, procedures, strategies) will students need in order to perform effectively and achieve desired results? What activities will equip students with the needed knowledge and skills? What will need to be taught and coached, and how should it best be taught, in light of performance goals? What materials and resources are best suited to accomplish these goals?

Note that the specifics of instructional planning—choices about teaching methods, sequence of lessons, and resource materials—can be successfully completed only after we identify desired results and assessments and consider what they imply. Teaching is a means to an end. Having a clear goal helps to focus our planning and guide purposeful action toward the intended results.

MISCONCEPTION ALERT!

When we speak of evidence of desired results, we are referring to evidence gathered through a variety of formal and informal assessments during a unit of study or a course. We are not alluding only to end-of-teaching tests or culminating tasks. Rather, the collected evidence we seek may well include traditional quizzes and tests, performance tasks and projects, observations and dialogues, as well as students' self-assessments gathered over time.

Backward design may be thought of, in other words, as purposeful task analysis: Given a worthy task to be accomplished, how do we best get everyone equipped? Or we might think of it as building a wise itinerary, using a map: Given a destination, what's the most effective and efficient route? Or we might think of it as planning for coaching, as suggested earlier: What must learners master if they are to effectively perform? What will count as evidence *on the field,* not merely in drills, that they really get it and are ready to *perform with understanding, knowledge, and skill* on their own? How will the learning be designed so that learners' capacities are developed through use and feedback?

This is all quite logical when you come to understand it, but "backward" from the perspective of much habit and tradition in our field. A major change from common practice occurs as designers must begin to think about assessment *before* deciding what and how they will teach. Rather than creating assessments near the conclusion of a unit of study (or relying on the tests provided by textbook publishers, which may not completely or appropriately assess our standards and goals), backward design calls for us to make our goals or standards specific and concrete, in terms of assessment evidence, as we begin to plan a unit or course.

The logic of backward design applies regardless of the learning goals. For example, when starting from a state content standard, curriculum designers need to determine the appropriate assessment evidence stated or implied in the standard. Likewise, a staff developer should determine what evidence will indicate that the adults have learned the intended knowledge or skill before planning the various workshop activities.

The rubber meets the road with assessment. Three different teachers may all be working toward the same content standards, but if their assessments vary considerably, how are we to know which students have achieved what? Agreement on needed evidence of learning leads to greater curricular coherence and more reliable evaluation by teachers. Equally important is the long-term gain in teacher, student, and parent insight about what does and does not count as evidence of meeting complex standards.

This view of focusing intently on the desired learning is hardly radical or new. Tyler (1949) described the logic of backward design clearly and succinctly more than 50 years ago:

> *Educational objectives become the criteria by which materials are selected, content is outlined, instructional procedures are developed, and tests and examinations are prepared. . . .*
>
> *The purpose of a statement of objectives is to indicate the kinds of changes in the student to be brought about so that instructional activities can be planned and developed in a way likely to attain these objectives. (pp. 1, 45)*

And in his famous book, *How to Solve It,* originally published in 1945, Polya specifically discusses "thinking backward" as a strategy in problem solving going back to the Greeks:

> *There is a certain psychological difficulty in turning around, in going away from the goal, in working backwards. . . . Yet, it does not take a genius to solve a concrete problem working backwards; anyone can do it with a little common sense. We concentrate on the desired end, we visualize the final position in which we would like to be. From what foregoing position could we get there? (p. 230)*

These remarks are old. What is perhaps new is that we offer herein a helpful process, a template, a set of tools, and design standards to make the plan and resultant student performance more likely to be successful by design than by good fortune. As a 4th grade teacher from Alberta, Canada, put it, "Once I had a way of clearly defining the end in mind, the rest of the unit 'fell into place.'"

The twin sins of activity-based and coverage-based design reflect a failure to think through purpose in this backward-design way. With this in mind, let's revisit the two fictitious vignettes from the Introduction. In the apples vignette, the unit seems to focus on a particular theme (harvest time), through a specific and familiar object (apples). But as the depiction reveals, the unit has no real depth because there is no enduring learning for the students to derive. The work is *hands-on* without being *minds-on,* because students do not need to (and are not really challenged to) extract sophisticated ideas or connections. They don't have to work at understanding; they need only engage in the activity. (Alas, it is common to reward students for mere engagement as opposed to understanding; engagement is necessary, but not sufficient, as an end result.)

Moreover, when you examine the apples unit it becomes clear that it has no overt priorities—the activities appear to be of equal value. The students' role is merely to participate in mostly enjoyable activities, without having to demonstrate that they understand any big ideas at the core of the subject (excuse the pun). All activity-based—as opposed to results-based—teaching shares the weakness of the apples unit: Little in the design asks students to derive intellectual fruit from the unit (sorry!). One might characterize this activity-oriented approach as "faith in learning by osmosis." Is it likely that individual students will learn a few interesting things about apples? Of course. But, in the absence of a learning plan with clear goals, how likely is it that students will develop shared understandings on which future lessons might build? Not very.

In the world history vignette, the teacher covers vast amounts of content during the last quarter of the year. However, in his harried march to get through a textbook, he apparently does not consider what the students will understand and apply from the material. What kind of intellectual scaffolding is provided to guide students through the important ideas? How are students expected to use those ideas to make meaning of the many facts? What performance goals would help students know how to take notes for maximal effective use by the course's end? Coverage-based instruction amounts to the teacher merely talking, checking off topics, and moving on, irrespective of whether students understand or are confused. This approach might be termed "teaching by mentioning it." Coverage-oriented teaching typically relies on a textbook, allowing it to define the content and sequence of instruction. In contrast, we propose that results-oriented teaching employ the textbook as a resource but not the syllabus.

A backward design template

Having described the backward design process, we now put it together in a useful format—a template for teachers to use in the design of units that focus on understanding.

Many educators have observed that backward design is common sense. Yet when they first start to apply it, they discover that it feels unnatural. Working this way may seem a bit awkward and time-consuming until you get the hang of it. But the effort is worth it—just as the learning curve on good software is worth it. We think of Understanding by Design as software, in fact: a set of tools for making you ultimately more productive. Thus, a practical cornerstone of Understanding by Design is a design template that is meant to reinforce the appropriate habits of mind needed to complete designs for student understanding and to avoid the habits that are at the heart of the twin sins of activity-based and coverage-based design.

Figure 1.2 provides a preliminary look at the UbD Template in the form of a one-page version with key planning questions included in the various fields. This format guides the teacher to the various UbD elements while visually conveying the idea of backward design. Later chapters present a more complete account of the template and each of its fields.

Although this one-page version of the template does not allow for great detail, it has several virtues. First, it provides a *gestalt,* an overall view of backward design, without appearing overwhelming. Second, it enables a quick check of alignment—the extent to which the assessments (Stage 2) and learning activities (Stage 3) align with identified goals (Stage 1). Third, the template

Figure 1.2
1-Page Template with Design Questions for Teachers

Stage 1—Desired Results

Established Goals: G

- What relevant goals (e.g., content standards, course or program objectives, learning outcomes) will this design address?

Understandings: U

Students will understand that . . .

- What are the big ideas?
- What specific understandings about them are desired?
- What misunderstandings are predictable?

Essential Questions: Q

- What provocative questions will foster inquiry, understanding, and transfer of learning?

Students will know . . . K

- What key knowledge and skills will students acquire as a result of this unit?
- What should they eventually be able to do as a result of such knowledge and skills?

Students will be able to . . . S

Stage 2—Assessment Evidence

Performance Tasks: T

- Through what authentic performance tasks will students demonstrate the desired understandings?
- By what criteria will performances of understanding be judged?

Other Evidence: OE

- Through what other evidence (e.g., quizzes, tests, academic prompts, observations, homework, journals) will students demonstrate achievement of the desired results?
- How will students reflect upon and self-assess their learning?

Stage 3—Learning Plan

Learning Activities: L

What learning experiences and instruction will enable students to achieve the desired results? How will the design

W = Help the students know Where the unit is going and What is expected? Help the teacher know Where the students are coming from (prior knowledge, interests)?
H = Hook all students and Hold their interest?
E = Equip students, help them Experience the key ideas and Explore the issues?
R = Provide opportunities to Rethink and Revise their understandings and work?
E = Allow students to Evaluate their work and its implications?
T = Be Tailored (personalized) to the different needs, interests, and abilities of learners?
O = Be Organized to maximize initial and sustained engagement as well as effective learning?

can be used to review existing units that teachers or districts have developed. Finally, the one-page template provides an initial design frame. We also have a multipage version that allows for more detailed planning, including, for example, a Performance Task Blueprint and a day-by-day calendar for listing and sequencing key learning events. The *Understanding by Design Professional Development Workbook* (McTighe & Wiggins, 2004, pp. 46–51) includes a six-page template that allows for more detailed planning.

We regularly observe that teachers begin to internalize the backward design process as they work with the UbD Template. Stage 1 asks designers to consider what they want students to understand and then to frame those understandings in terms of questions. In completing the top two sections of the Stage 1 portion of the template, users are prompted to identify the Understandings and Essential Questions to establish a larger context into which a particular unit is nested.

Stage 2 prompts the designer to consider a variety of assessment methods for gathering evidence of the desired Understandings. The two-box graphic organizer then provides spaces for specifying the particular assessments to be used during the unit. Designers need to think in terms of collected evidence, not a single test or performance task.

Stage 3 calls for a listing of the major learning activities and lessons. When it is filled in, the designer (and others) should be able to discern what we call the "WHERETO" elements.

The *form* of the template offers a means to succinctly present the design unit; its *function* is to guide the design process. When completed, the template can be used for self-assessment, peer review, and sharing of the completed unit design with others.

To better understand the template's benefits for the teacher-designer, let's take a look at a completed template. Figure 1.3 shows a completed three-page version of the template for a unit on nutrition.

Notice that the template in Figure 1.3 supports backward design thinking by making the longer-term goals far more explicit than is typical in lesson planning, and we can follow those goals through Stages 2 and 3 to ensure that the design is coherent. The focus on big ideas in Stage 1 is transparent, without sacrificing the more discrete elements of knowledge and skill. Finally, by calling for appropriately different types of assessment, the template reminds us that we typically need varied evidence and assessments grounded in performance to show transfer, if understanding is our aim.

Design standards

Accompanying the UbD Template is a set of Design Standards corresponding to each stage of backward design. The standards offer criteria to use during development and for quality control of completed unit designs. Framed as questions, the UbD Design Standards serve curriculum designers in the same

Figure 1.3
3-Page Nutrition Example

Stage 1—Identify Desired Results

Established Goals:

G

Standard 6—Students will understand essential concepts about nutrition and diet.
6a—Students will use an understanding of nutrition to plan appropriate diets for themselves and others.
6c—Students will understand their own individual eating patterns and ways in which those patterns may be improved.

What essential questions will be considered?

Q

- What is healthful eating?
- Are you a heathful eater? How would you know?
- How could a healthy diet for one person be unhealthy for another?
- Why are there so many health problems in the United States caused by poor nutrition despite all the available information?

What understandings are desired?

U

Students will understand that . . .

- A balanced diet contributes to physical and mental health.
- The USDA food pyramid presents *relative* guidelines for nutrition.
- Dietary requirements vary for individuals based on age, activity level, weight, and overall health.
- Healthful living requires an individual to act on available information about good nutrition even if it means breaking comfortable habits.

What key knowledge and skills will students acquire as a result of this unit?

K

Students will know . . .

- Key terms—protein, fat, calorie, carbohydrate, cholesterol.
- Types of foods in each food group and their nutritional values.
- The USDA food pyramid guidelines.
- Variables influencing nutritional needs.
- General health problems caused by poor nutrition.

S

Students will be able to . . .

- Read and interpret nutrition information on food labels.
- Analyze diets for nutritional value.
- Plan balanced diets for themselves and others.

Figure 1.3 (continued)

3-Page Nutrition Example

Stage 2—Determine Acceptable Evidence

What evidence will show that students understand?

Performance Tasks:

T

You Are What You Eat—Students create an illustrated brochure to teach younger children about the importance of good nutrition for healthful living. They offer younger students ideas for breaking bad eating habits.

Chow Down—Students develop a three-day menu for meals and snacks for an upcoming Outdoor Education camp experience. They write a letter to the camp director to explain why their menu should be selected (by showing that it meets the USDA food pyramid recommendations, yet it is tasty enough for the students). They include at least one modification for a specific dietary condition (diabetic or vegetarian) or religious consideration.

What other evidence needs to be collected in light of Stage 1 Desired Results?

Other Evidence:
(e.g., tests, quizzes, prompts, work samples, observations)

OE

Quiz—The food groups and the USDA food pyramid

Prompt—Describe two health problems that could arise as a result of poor nutrition and explain how these could be avoided.

Skill Check—Interpret nutritional information on food labels.

Student Self-Assessment and Reflection:

SA

1. Self-assess the brochure, *You Are What You Eat.*
2. Self-assess the camp menu, *Chow Down.*
3. Reflect on the extent to which you eat healthfully at the end of unit (compared with the beginning).

Figure 1.3 (continued)

3-Page Nutrition Example

Stage 3—Plan Learning Experiences

What sequence of teaching and learning experiences will equip students to engage with, develop, and demonstrate the desired understandings? Use the following sheet to list the key teaching and learning activities in sequence. Code each entry with the appropriate initials of the WHERETO elements.

1. Begin with an entry question (Can the foods you eat cause zits?) to hook students into considering the effects of nutrition on their lives. H
2. Introduce the Essential Questions and discuss the culminating unit performance tasks (Chow Down and Eating Action Plan). W
3. Note: Key vocabulary terms are introduced as needed by the various learning activities and performance tasks. Students read and discuss relevant selections from the Health textbook to support the learning activities and tasks. As an ongoing activity, students keep a chart of their daily eating and drinking for later review and evaluation. E
4. Present concept attainment lesson on the food groups. Then have students practice categorizing pictures of foods accordingly. E
5. Introduce the Food Pyramid and identify foods in each group. Students work in groups to develop a poster of the Food Pyramid containing cut-out pictures of foods in each group. Display the posters in the classroom or hallway. E
6. Give quiz on the food groups and Food Pyramid (matching format). E
7. Review and discuss the nutrition brochure from the USDA. Discussion question: Must everyone follow the same diet to be healthy? R
8. Working in cooperative groups, students analyze a hypothetical family's diet (deliberately unbalanced) and make recommendations for improved nutrition. Teacher observes and coaches students as they work. E-2
9. Have groups share their diet analyses and discuss as a class. E, E-2 (Note: Teacher collects and reviews the diet analyses to look for misunderstandings needing instructional attention.)
10. Each student designs an illustrated nutrition brochure to teach younger children about the importance of good nutrition for healthy living and the problems associated with poor eating. This activity is completed outside of class. E, T
11. Students exchange brochures with members of their group for a peer assessment based on a criteria list. Allow students to make revisions based on feedback. R, E-2
12. Show and discuss the video, "Nutrition and You." Discuss the health problems linked to poor eating. E
13. Students listen to, and question, a guest speaker (nutritionist from the local hospital) about health problems caused by poor nutrition. E
14. Students respond to written prompt: Describe two health problems that could arise as a result of poor nutrition and explain what changes in eating could help to avoid them. (These are collected and graded by teacher.) E-2
15. Teacher models how to read and interpret food label information on nutritional values. Then have students practice using donated boxes, cans, and bottles (empty!). E
16. Students work independently to develop the three-day camp menu. Evaluate and give feedback on the camp menu project. Students self- and peer-assess their projects using rubrics. E-2, T
17. At the conclusion of the unit, students review their completed daily eating chart and self-assess the healthfulness of their eating. Have they noticed changes? Improvements? Do they notice changes in how they feel and their appearance? E-2
18. Students develop a personal "eating action plan" for healthful eating. These are saved and presented at upcoming student-involved parent conferences. E-2, T
19. Conclude the unit with student self-evaluation regarding their personal eating habits. Have each student develop a personal action plan for their "healthful eating" goal. E-2, T

way that a scoring rubric serves students. When presented to students before they begin their work, the rubric provides them with a performance target by identifying the important qualities toward which they should strive. Similarly, the Design Standards specify the qualities of effective units according to the Understanding by Design framework. Figure 1.4 (p. 28) presents the four UbD Design Standards with accompanying indicators.

The standards contribute to design work in three ways:

- *As a reference point during design*—Teachers can periodically check to see, for example, if the identified understandings are truly big and enduring, or if the assessment evidence is sufficient. Like a rubric, the questions serve as reminders of important design elements to include, such as a focus on Essential Questions.
- *For use in self-assessment and peer reviews of draft designs*—Teachers and peers can use the criteria to examine their draft units to identify needed refinements, such as using the facets to dig deeper into an abstract idea.
- *For quality control of completed designs*—The standards can then be applied by independent reviewers (e.g., curriculum committees) to validate the designs before their distribution to other teachers.

Our profession rarely subjects teacher-designed units and assessments to this level of critical review. Nonetheless, we have found structured peer reviews, guided by design standards, to be enormously beneficial—both to teachers and their designs (Wiggins, 1996, 1997). Participants in peer review sessions regularly comment on the value of sharing and discussing curriculum and assessment designs with colleagues. We believe that such sessions are a powerful approach to professional development, because the conversations focus on the heart of teaching and learning.

We cannot stress enough the importance of using design standards to regularly review curriculum—existing units and courses as well as new ones being developed. It is often difficult for educators, both novice and veteran, to get in the habit of self-assessing their designs against appropriate criteria. A prevailing norm in our profession seems to be, "If I work hard on planning, it must be good." The UbD Design Standards help to break that norm by providing a means for quality control. They help us validate our curriculum's strengths, while revealing aspects that need improvement.

In addition to using the UbD Design Standards for self-assessment, the quality of the curriculum product (unit plan, performance assessment, course design) is invariably enhanced when teachers participate in a structured peer review in which they examine one another's unit designs and share feedback and suggestions for improvement. Such "critical friend" reviews provide feedback to designers, help teachers internalize the qualities of good design, and offer opportunities to see alternate design models. ("Gee, I never thought about beginning a unit with a problem. I think I'll try that in my next unit.")

Figure 1.4
UbD Design Standards

Stage 1—To what extent does the design focus on the big ideas of targeted content?

Consider: Are . . .

- ❍ The targeted understandings enduring, based on transferable, big ideas at the heart of the discipline and in need of uncoverage?
- ❍ The targeted understandings framed by questions that spark meaningful connections, provoke genuine inquiry and deep thought, and encourage transfer?
- ❍ The essential questions provocative, arguable, and likely to generate inquiry around the central ideas (rather than a "pat" answer)?
- ❍ Appropriate goals (e.g., content standards, benchmarks, curriculum objectives) identified?
- ❍ Valid and unit-relevant knowledge and skills identified?

Stage 2—To what extent do the assessments provide fair, valid, reliable, and sufficient measures of the desired results?

Consider: Are . . .

- ❍ Students asked to exhibit their understanding through authentic performance tasks?
- ❍ Appropriate criterion-based scoring tools used to evaluate student products and performances?
- ❍ Various appropriate assessment formats used to provide additional evidence of learning?
- ❍ The assessments used as feedback for students and teachers, as well as for evaluation?
- ❍ Students encouraged to self-assess?

Stage 3—To what extent is the learning plan effective and engaging?

Consider: Will the students . . .

- ❍ Know where they're going (the learning goals), why the material is important (reason for learning the content), and what is required of them (unit goal, performance requirements, and evaluative criteria)?
- ❍ Be hooked—engaged in digging into the big ideas (e.g., through inquiry, research, problem solving, and experimentation)?
- ❍ Have adequate opportunities to explore and experience big ideas and receive instruction to equip them for the required performances?
- ❍ Have sufficient opportunities to rethink, rehearse, revise, and refine their work based upon timely feedback?
- ❍ Have an opportunity to evaluate their work, reflect on their learning, and set goals?

Consider: Is the learning plan . . .

- ❍ Tailored and flexible to address the interests and learning styles of all students?
- ❍ Organized and sequenced to maximize engagement and effectiveness?

Overall Design—To what extent is the entire unit coherent, with the elements of all three stages aligned?

Design tools

In addition to the design standards, we have developed and refined a comprehensive set of design tools to support teachers and curriculum developers. This is hard work! We have found that an array of scaffolds—prompts, organizers, idea sheets, and examples—help educators produce higher-quality designs. A full set of these resources is available in the *UbD Professional Development Workbook.*

We think that a good template serves as an intelligent tool. It provides more than a place to write in ideas. It focuses and guides the designer's thinking throughout the design process to make high-quality work more likely. In practice, curriculum designers work from a copy of the template, supported by specific design tools and numerous filled-in examples of good unit designs. In this way, we practice what we preach with students; models and design standards are provided up front to focus designer performance from the start.[1]

But why do we refer to the template, design standards, and corresponding design tools as "intelligent"? Just as a physical tool (e.g., a telescope, an automobile, or a hearing aid) extends human capabilities, an intelligent tool enhances performance on cognitive tasks, such as the design of learning units. For example, an effective graphic organizer, such as a story map, helps students internalize the elements of a story in ways that enhance their reading and writing of stories. Likewise, by routinely using the template and design tools, users will likely develop a mental template of the key ideas presented in this book: the logic of backward design, thinking like an assessor, the facets of understanding, WHERETO, and design standards.

By embodying the Understanding by Design elements in tangible forms (i.e., the template and design tools), we seek to support educators in learning and applying these ideas. Thus, the design tools are like training wheels, providing a steadying influence during those periods of disequilibrium brought on by new ideas that may challenge established and comfortable habits. Once the key ideas of Understanding by Design are internalized, however, and regularly applied, the explicit use of the tools becomes unnecessary, just as the young bicycle rider sheds the training wheels after achieving balance and confidence.

> **MISCONCEPTION ALERT!**
>
> Though the three stages present a logic of design, it does not follow that this is a step-by-step process in actuality. As we argue in Chapter 11, don't confuse the logic of the final product with the messy process of design work. It doesn't matter exactly where you start or how you proceed, *as long as you end up with a coherent design* reflecting the logic of the three stages. The final outline of a smoothly flowing college lecture rarely reflects the back-and-forth (iterative) thought process that went into its creation.

Backward design in action with Bob James

Setting: *We are inside the head of Bob James, a 6th grade teacher at Newtown Middle School, as he begins to design a three-week unit on nutrition. His ultimate* *design will be the unit provided above in Figure 1.3. But Bob is new to UbD, so his design will unfold and be revised over time. Throughout the book we'll show his thinking—and rethinking—as he considers the full meaning of the template elements.*

Stage 1: Identify desired results

The template asks me to highlight the goals of the unit, and for me that means drawing upon our state standards. In reviewing our standards in health, I found three content standards on nutrition that are benchmarked to this age level:

- Students will understand essential concepts about nutrition.
- Students will understand elements of a balanced diet.
- Students will understand their own eating patterns and ways in which these patterns may be improved.

Using these standards as the starting point, I need to decide what I want my students to take away from the unit. Knowledge and skill are what I have always focused on: knowledge of the food pyramid, the ability to read labels in the store and at home, and so on. Although I've never deliberately thought about *understandings*, per se, I like the concept and think that it will help me focus my teaching and limited class time on the truly important aspects of this unit.

As I think about it, I guess what I'm really after has something to do with an understanding of the elements of good nutrition so students can plan a balanced diet for themselves and others. The big ideas have to do with nutrition and planning meals in a feasible way. Then, the important questions are, So, what is good for you? What isn't? How do you know? What makes it difficult to know and to eat right? (The good taste of junk food makes it difficult!)

This idea is clearly important, because planning nutritious menus is an authentic, lifelong need and a way to apply this knowledge. I'm still a little unclear about what "an understanding" means, though, in this context. I'll need to reflect further on what an understanding is and how it goes beyond specific knowledge and its use. The basic concepts of nutrition are fairly straightforward, after all, as are the skills of menu planning. Does anything in the unit require, then, any in-depth and deliberate *uncoverage*? Are there typical misunderstandings, for example, that I should more deliberately focus on?

Well, as I think about it, I have found that many students harbor the two misconceptions that if food is good for you, it must taste bad; and if it is sold in famous and popular places, it must be okay. One of my goals in this unit is to dispel these myths so that the students won't have an automatic aversion to healthy food and unwittingly eat too much unhealthy stuff. In terms of the potential for engagement—no problem there. Anything having to do with food is a winner with 10- and 11-year-olds. And there are some points to menu planning (such as balancing cost, variety, taste, and dietary needs) that are not at all obvious. This way of thinking about the unit will enable me to better focus on these points.

Stage 2: Determine acceptable evidence

This will be a bit of a stretch for me. Typically in a three- or four-week unit like this one, I give one or two quizzes; have a project, which I grade; and conclude with a unit test (generally multiple choice or matching). Even though this approach to assessment makes grading and justifying the grades fairly easy, I have always felt a bit uneasy that these assessments don't reflect the point of the unit and that the project grade sometimes has less to do with the key ideas and more to do with effort. I think I tend to test what is easy to test instead of assessing for my deeper goals, above and beyond nutritional facts. In fact, one thing that has always disturbed me is that the kids tend to focus on their grades rather than on their learning. Perhaps the way I've used the assessments—more for grading purposes than to help shape and document learning—has contributed somewhat to their attitude.

Now I need to think about what would serve as evidence of the ideas I'm focusing on. After reviewing some examples of performance tasks and discussing "application" ideas with my colleagues, I have decided tentatively on the following task:

> *Because we have been learning about nutrition, the camp director at the outdoor education center has asked us to propose a nutritionally balanced menu for our three-day trip to the center later this year. Using the food pyramid guidelines and the nutrition facts on food labels, design a plan for three days, including three meals and three snacks (a.m., p.m., and campfire). Your goal: a tasty and nutritionally balanced menu.*

I'm excited about this idea because it asks students to demonstrate what I really want them to take away from the unit. This task also links well with one of our unit projects: to analyze a hypothetical family's diet for a week and propose ways to improve their nutrition. With this task and project in mind, I can now use my quizzes to check students' knowledge of the food groups and food pyramid recommendations, and a lengthier test to check for their understanding of how a nutritionally deficient diet contributes to health problems. Hey! This is one of the better assessment plans I have designed for a unit, and I think that the task will motivate students as well as provide evidence of their understanding.

Stage 3: Plan learning experiences and instruction

This is my favorite part of planning—deciding what activities the students will do during the unit and what resources and materials we'll need for those activities. But according to what I'm learning about backward design, I'll need to think first about what essential knowledge and skills my students will need if they're going to be able to demonstrate in performance the understandings I'm after.

Well, they'll need to know about the different food groups and the types of foods found in each group so that they'll understand the USDA food pyramid recommendations. They'll also need to know about human nutritional needs for carbohydrates, protein, sugar, fat, salt, vitamins, and minerals, and about the various foods that provide them. They'll have to learn about the minimum daily requirements for these nutritional elements and about various health problems that arise from poor nutrition. In terms of skills, they'll have to learn how to read and interpret the nutrition-fact labels on foods and how to scale a recipe up or down, because these skills are necessary for their culminating project—planning healthy menus for camp.

Now for the learning experiences. I'll use resources that I've collected during the past several years—a pamphlet from the USDA on the food groups and the food pyramid recommendations; a wonderful video, "Nutrition for You"; and, of course, our health textbook (which I now plan to use selectively). As I have for the past three years, I'll invite the nutritionist from the local hospital to talk about diet and health and how to plan healthy menus. I've noticed that the kids really pay attention to a real-life user of information they're learning.

My teaching methods will follow my basic pattern—a blend of direct instruction, inductive methods, cooperative-learning group work, and individual activities.

Planning backward to produce this new draft has been helpful. I now can more clearly see and state what knowledge and skills are essential, given my goals for the unit. I'll be able to concentrate on the more important aspects of the topic (and relieve some guilt that I'm not covering everything). It's also interesting to realize that even though some sections of the textbook chapters on nutrition will be especially useful (for instance, the descriptions of health problems arising from poor nutrition), other sections are not as informative as other resources I'll now use (the brochure and the video). In terms of assessment, I now know more clearly what I need to assess using traditional quizzes and tests, and why the performance task and project are needed—to have students demonstrate their understanding. I'm getting a feel for backward design.

Comments on the design process

Notice that the process of developing this draft nutrition unit reveals four key aspects of backward design:

1. The assessments—the performance tasks and related sources of evidence—are thought through prior to the lessons being fully developed. The assessments serve as teaching targets for sharpening the focus of instruction and editing the past lesson plans, because they define in very specific terms what we want students to understand and be able to do. The teaching is then thought of as *enabling* performance. These assessments also guide decisions about what content needs to be emphasized versus that which is not really essential.

2. It is likely that familiar and favorite activities and projects will have to be further modified in light of the evidence needed for assessing targeted standards. For instance, if the apples unit described in the Introduction were planned using this backward design process, we would expect to see revisions in some of the activities to better support the desired results.

3. The teaching methods and resource materials are chosen last, with the teacher keeping in mind the work that students must produce to meet the standards. For example, rather than focusing on cooperative learning because it's a popular strategy, the question from a backward-design perspective becomes, What instructional strategies will be most effective in helping us reach our targets? Cooperative learning may or may not be the best approach, given the particular students and standards.

4. The role of the textbook may shift from being the primary resource to being a support. Indeed, the 6th grade teacher planning the nutrition unit realized the limitations of relying on the text if he is to meet his goals. Given other valuable resources (the USDA materials, the video, and the nutritionist), he no longer felt compelled to cover the book word for word.

This introductory look is intended to present a preliminary sketch of the big picture of a design approach. Bob James will be refining his unit plan (and changing his thinking a few times) as he gains greater insight into understanding, essential questions, valid assessment, and the related learning activities.

A preview

Figure 1.5 presents the key elements of the UbD approach and thus an outline of points to come in the book. In the following chapters we "uncover" this design process, examining its implications for the development and use of assessments, the planning and organization of curriculum, and the selection of powerful methods of teaching. But a few explanatory points about each column in Figure 1.5 are appropriate to prepare you for what is to come throughout the book.

The chart is best read from left to right, one row at a time, to see how the three stages of design might look in practice. An outline of the three-stage design process for each of the three basic elements (the desired results, the assessment evidence, and the learning plan) is highlighted in the column headings. Begin with a key design question; ponder how to narrow the possibilities through intelligent priorities (Design Considerations); self-assess, self-adjust, and finally critique each element of design against appropriate criteria (Filters); and end up with a product that meets appropriate design standards in light of the achievement target (What the Final Design Accomplishes).

In summary, backward design yields greater coherence among desired results, key performances, and teaching and learning experiences, resulting in better student performance—the purpose of design.

Figure 1.5
The UbD Design Matrix

Key Design Questions	Chapters of the Book	Design Considerations	Filters (Design Criteria)	What the Final Design Accomplishes
Stage 1 • What are worthy and appropriate results? • What are the key desired learnings? • What should students come away understanding, knowing, and able to do? • What big ideas can frame all these objectives?	• Chapter 3—Gaining Clarity on Our Goals • Chapter 4—The Six Facets of Understanding • Chapter 5—Essential Questions: Doorways to Understanding • Chapter 6—Crafting Understandings	• National standards • State standards • Local standards • Regional topic opportunities • Teacher expertise and interest	• Focused on big ideas and core challenges	• Unit framed around enduring understandings and essential questions, in relation to clear goals and standards
Stage 2 • What is evidence of the desired results? • In particular, what is appropriate evidence of the desired understanding?	• Chapter 7—Thinking like an Assessor • Chapter 8—Criteria and Validity	• Six facets of understanding • Continuum of assessment types	• Valid • Reliable • Sufficient	• Unit anchored in credible and useful evidence of the desired results
Stage 3 • What learning activities and teaching promote understanding, knowledge, skill, student interest, and excellence?	• Chapter 9—Planning for Learning • Chapter 10—Teaching for Understanding	• Research-based repertoire of learning and teaching strategies • Appropriate and enabling knowledge and skill	Engaging and effective, using the elements of WHERETO: • *Where* is it going? • *Hook* the students • *Explore* and *equip* • *Rethink* and *revise* • *Exhibit* and *evaluate* • *Tailor* to student needs, interests, and styles • *Organize* for maximum engagement and effectiveness	• Coherent learning activities and teaching that will evoke and develop the desired understandings, knowledge, and skill; promote interest; and make excellent performance more likely

Innovations in Teaching Undergraduate Biology and Why We Need Them

William B. Wood

Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, Colorado 80309-0347; email: wood@colorado.edu

Annu. Rev. Cell Dev. Biol. 2009. 25:93–112

First published online as a Review in Advance on July 7, 2009

The *Annual Review of Cell and Developmental Biology* is online at cellbio.annualreviews.org

This article's doi: 10.1146/annurev.cellbio.24.110707.175306

1081-0706/09/1110-0093$20.00

Key Words

active learning, concept inventories, course transformation, discipline-based educational research, formative assessment, pedagogy

Abstract

A growing revolution is under way in the teaching of introductory science to undergraduates. It is driven by concerns about American competitiveness as well as results from recent educational research, which explains why traditional teaching approaches in large classes fail to reach many students and provides a basis for designing improved methods of instruction. Discipline-based educational research in the life sciences and other areas has identified several innovative promising practices and demonstrated their effectiveness for increasing student learning. Their widespread adoption could have a major impact on the introductory training of biology students.

Contents

DEFINING THE CHALLENGE

Two principal forces are generating momentum for a revolution in the way biology and other sciences are taught in high schools, colleges, and universities (DeHaan 2005). First, there are deep concerns about American international competitiveness, amid indications that the U.S. is doing a relatively poor job at retaining and training students in the science, technology, engineering, and mathematics (STEM) disciplines (DoE 2000, NAS 2004). Too many talented students get the impression from introductory courses that science is simply a collection of facts to be memorized and consequently drop out of STEM majors with little understanding or appreciation of what science is all about (Seymour & Hewitt 1997). For students who do major in life sciences, there is concern that future research biologists are being inadequately trained (NRC 2003, AAMC-HHMI 2009).

STEM: science, technology, engineering, and mathematics

DBER: discipline-based educational research

The second driving force for reform is recent research from educators and cognitive scientists into how students learn. This research, discussed further below, provides strong evidence that the traditional teaching methods employed in most secondary-school and undergraduate introductory courses are far from optimal for promoting student learning. Alternative research-based teaching methods have been developed and shown to be more effective, and a small but growing number of informed STEM faculty and administrators are pushing for their adoption.

Beyond the general findings about how students learn, there is now a substantial body of discipline-based educational research (DBER) dealing with teaching and learning of specific STEM disciplines. This review refers to some of the more important general findings on how students learn, but it primarily highlights results and applications from recent DBER and, more specifically, life sciences education research. It focuses on teaching and learning for undergraduates, particularly in large courses, where innovation is most needed.

HISTORY AND CURRENT STATE OF DISCIPLINE-BASED EDUCATIONAL RESEARCH

DBER grew out of the efforts of physicists in the mid-1980s, who discovered that most undergraduate students in their introductory courses were gaining only very superficial knowledge from traditional methods of instruction (Halloun & Hestenes 1985, Hestenes et al. 1992). Rather than integrated conceptual understanding and creative problem solving, students were learning fragmented factual information and rote problem solving methods, while retaining many misconceptions about physical phenomena. To gain some measure of student understanding, physicists developed the Force Concept Inventory (FCI), a simple multiple-choice test of basic concepts and common misconceptions about Newtonian physics of everyday events written in simple language and requiring no sophisticated mathematics (Hestenes et al. 1992). By administering the FCI at the beginning and the end of an introductory course, instructors could obtain a measure of gains in student conceptual learning. They could then experiment with different instructional approaches and test them for efficacy. These physicists showed that adopting a small number of nontraditional promising practices in course design and implementation could substantially increase student learning gains. These practices, and their basis in more general educational research on how people learn, are described in the following sections.

After a lag of several years, instructors in other STEM disciplines began to make similar observations about their students and to initiate similar efforts at improving instruction. The empirical approach of varying instructional methods and measuring effects on student learning has been called "scientific teaching" (Handelsman et al. 2004, Wieman 2007). Many DBE researchers doing this work are practicing scientists trained in their disciplines who have also learned educational research methods and taken up DBER as a sideline. Some schools of education have added DBER practitioners trained as educators to their faculties. In addition, some university science departments, particularly in physics but increasingly in other STEM disciplines, now include staff or tenure-track DBE researchers (NAS 2005) and are beginning to offer graduate training and degrees in DBER.

FCI: force concept inventory

DBER is published in a variety of education journals, some general and some that are discipline-specific, sponsored by STEM professional societies. A few scientific journals, including *Nature*, *Science*, *PLoS Biology*, and *Genetics*, have also begun publishing DBER articles, generally in an education section. **Table 1** lists some of the more widely read general and discipline-specific educational journals that publish DBER in life sciences.

HOW STUDENTS LEARN

New ideas about teaching and learning began to receive public attention in the 1960s. Popular iconoclasts such as Holt (1964, 1967) and Kozol (1967), building on earlier ideas (Dewey 1916, Ausubel 1963), pointed out the shortcomings of passive learning environments for learners of all ages and advocated instead more student-centered, open classrooms that promoted active learning through hands-on experience, by doing rather than by simply listening, reading, and watching. These writers, considered radicals in their time, articulated ideas about optimal conditions for meaningful learning that have since been tested and validated by a large body of educational research. Also during the past three decades, advances in cognitive science have begun to elucidate the neural activities and synaptic changes that accompany learning. Results of research in both education and cognition were reviewed in the seminal National Research Council (NRC) report *How People Learn: Brain, Mind, Experience, and School* (NRC 1999). The

Table 1 A partial list of journals that publish life sciences education research[a]

General scientific journals
Genetics
Nature
Science
PLoS Biology[b]
Education journals (sponsored by professional societies)
Advances in Physiology Education, 2001- (Amer. Physiol. Soc.)*†
Biochemistry and Molecular Biology Education, 2006- (Amer. Soc. Biochem. and Mol. Biol.)†
CBE-Life Sciences Education, 2002- (Am. Soc. Cell Biol.)[b,c]
Journal of Biological Education, 1990- (Brit. Inst. Biol.)[b]
Microbiology Education Journal (Amer. Soc. Microbiol.)[c]
Frontiers in Ecology and the Environment (Ecol. Soc. Am.)
General education journals
American Biology Teacher (Natl. Assoc. Biol. Teach.)[b,c]
Bioscene: Journal of College Biology Teaching[b]
BioScience (Am. Inst. Biol. Sci.)
International Journal of the Scholarship of Teaching and Learning[b,c]
Journal of College Science Teaching (Natl. Sci. Teach. Assoc.)

[a]For additional journal listings, see Dolan (2007).
[b]Open access.
[c]Higher standards: Research articles require assessment and outcome evidence for efficacy of a new course or intervention rather than simple descriptions of practice.

Constructivist: the view that individual learners must build their own knowledge structures, from experience and instruction, on a foundation of prior knowledge

Formative assessment: frequent, ongoing testing, usually during class, with the goal of monitoring understanding and providing feedback rather than judging performance

Summative assessment: high-stakes testing at the end of an instructional unit or course to judge student performance, e.g., mid-term and final exams

major conclusions from this research can be summarized as follows:

- Learning involves the elaboration of knowledge structures in long-term memory. According to this constructivist view of education (Dewey 1916; Ausubel 1963, 2000), effective instruction must begin at the level of a student's prior knowledge (which may include misconceptions). New information unrelated to prior knowledge is difficult to learn and remember.
- No two learners are the same: Learners differ in previous experience, previous instruction, preferred styles of learning, family background, cultural background, and so on. Diversity is an asset for collaborative work because different members of a group bring different perspectives and skills to bear, but it can hamper learning for some students unless the level and mode of instruction are appropriate for all.
- Learning is promoted by frequent feedback, that is, ongoing testing of new knowledge as students are acquiring it. Educators call this formative assessment, as opposed to summative assessment, which refers to high-stakes exams given after an extended period of instruction. Formative assessment provides valuable feedback to both instructor and students: Do students understand the concept just presented or discussed? Can they transfer this understanding to apply the concept in a new situation?
- Effective learning requires awareness and questioning of one's own learning process: How well do I understand this? What information do I need to understand it better? What do I not understand yet? Do I understand it well enough to transfer it, that is, apply it to a new situation? Educators call this awareness metacognition.

- Learning is enhanced in a community of learners who value the knowledge that is being learned. In early childhood this community is the family; at the university it could be a group of students working together to solve a problem or complete a research project.
- Learning changes the structure of the brain, and the extent of change increases with the degree of complexity, stimulation, and emotional involvement in the learning environment (Zull 2002). Active learning, in which a student's levels of motivation, curiosity, and attention are high, for example during a group effort to solve an intriguing problem, will be better retained than learning from relatively passive activities such as reading a text or listening to a lecture.
- Learning in a particular area of knowledge such as life sciences can be viewed as a continuum from novice to expert status, along which we would like to help our students progress. The knowledge of an expert constitutes a coherent structure into which new concepts can easily fit and from which relevant information can be efficiently retrieved. In contrast, new knowledge for the novice often appears to be a collection of unrelated facts, which are difficult to memorize and retain. In other words, experts see and make use of meaningful patterns and relationships in the information they possess, whereas novices cannot.

APPLICATION TO THE COLLEGE CLASSROOM

These general conclusions apply to teaching and learning of STEM disciplines at the undergraduate level:

- Effective instruction must build on students' prior knowledge (which may include misconceptions that require correction).
- Instructors should be aware of the student diversity in their classrooms and use a variety of teaching modes to optimize learning for all students.
- Classes should include frequent formative assessment to provide feedback to both instructors and students.
- Students should be encouraged to examine and monitor their own understanding of new concepts, for example, by explaining these concepts to their peers.
- Students should be encouraged to work cooperatively and collaboratively in small groups.
- In order to bring about the neurological changes that constitute learning, students should spend time actively engaged with the subject matter, for example, discussing, diagramming, solving problems, working on a research project, etc., in addition to or in place of listening passively to a lecture, reading the textbook, or consulting Web sites.

Most undergraduate college STEM classes, particularly in large introductory courses, are not designed around these principles, and it can be argued that this is one reason for the high attrition rates and generally superficial learning among introductory students in STEM disciplines. Educators have shown that effective instruction requires not only disciplinary content knowledge, for example, expertise in life sciences, but also pedagogical content knowledge, that is, understanding of and ability to apply known educational principles. Because graduate and postdoctoral training in STEM disciplines seldom includes any instruction in pedagogical practice, most university faculty are unaware of new knowledge about learning that could make their teaching more effective. Therefore, they simply teach the way they were taught in large classes, by traditional lecturing. We need to improve the way we teach undergraduates. The remainder of this article discusses evidence that applying the above principles to college classrooms can make a difference in how much and how well our students learn.

Transfer: application of knowledge learned in one context to a problem in a different context

Metacognition: the process of monitoring one's own learning process and level of understanding

EVIDENCE THAT RESEARCH-BASED TEACHING AT THE COLLEGE LEVEL INCREASES STUDENT LEARNING

Our best undergraduates, sometimes with little help from faculty, develop learning skills that incorporate the above principles, allowing them to progress toward expert knowledge regardless of how we teach them. However, many students, for whom studying means highlighting phrases in their textbooks and memorizing disconnected facts, fail to develop effective learning skills and consequently learn very little. Is there evidence that changes in teaching practices at the college level can significantly enhance student learning?

Physicists were the first to obtain such evidence, using the Force Concept Inventory (FCI; Hestenes et al. 1992) described above. The FCI became nationally accepted among physics instructors during the 1990s as a way to gauge student learning of Newtonian mechanics. Administering the FCI as a pre-test at the start of a course and then again as a post-test at the end yielded a raw learning gain for each student. For comparison of students with different levels of incoming knowledge, each raw gain was divided by the maximum possible gain for that student to arrive at a percentage normalized gain: $<g> = 100$(post-test score – pre-test score)/(100 – pre-test score).

In attempts to increase the generally low normalized gains seen in traditional introductory courses, physics education researchers transformed their courses with new teaching approaches following the principles described above: more class time devoted to active learning, more group problem solving, frequent formative assessment, and so on. They carried out controlled studies, for example, the same instructor teaching the same syllabus through traditional lectures in one semester and then using the new approaches in the following semester (e.g., Beichner 2008). Study after study indicated that students in the transformed courses substantially outperformed those in traditional courses. In a compelling landmark meta-analysis combining data from many such studies, Hake (1998) showed that for a sample of over 6000 students in 55 introductory physics courses nationwide, the average learning gains were nearly twice as high in transformed courses as in traditional courses.

Other STEM disciplines have lacked widely accepted assessment instruments comparable to the FCI until recently (see below). Nevertheless, several studies using some form of pre- and post-testing have also yielded results showing the greater efficacy of transformed courses. In the life sciences, an early study from the University of Oregon showed that students in the traditional introductory course learned substantially less than students in a workshop biology course, in which lecturing was almost entirely replaced by student group problem solving and other projects during class time (Udovic et al. 2002). Knight & Wood (2005) showed in a quasi-controlled study that even an incremental change, substituting 30–40% of lecturing during class time with more engaging student-centered activities (described below), led to increases in normalized learning gains averaging about 30% in a large upper-division developmental biology course. Similar results have been reported in large introductory biology courses (e.g., Smith et al. 2005, Armstrong et al. 2007, Freeman et al. 2007).

Clearly, concept inventories in life sciences would be valuable for continuation of this research (Garvin-Doxas et al. 2007) and several have recently been published for various subdisciplines including general biology (Klymkowsky et al. 2003), genetics (Bowling et al. 2008, Smith et al. 2008), and natural selection (Anderson et al. 2002). Libarkin (2008) has compiled a comprehensive current listing and comparison of concept inventories in STEM disciplines.

PROMISING PRACTICES FOR INCREASING STUDENT LEARNING

Many college faculty use Socratic dialog and student-centered group work in small classes and seminars, but they believe there is no alternative to lecturing when confronted with hundreds of students in an auditorium with fixed seats. However, innovative instructors pursuing DBER have developed and tested alternative teaching approaches that prove to be substantially more effective than traditional lectures. This research has identified several promising practices for transforming large classes in ways that enhance student learning and conceptual understanding (reviewed in Handelsman et al. 2007).

Froyd (2008) has introduced a useful rating of promising practices based on two criteria: (*a*) practicality of implementation (breadth of applicability to STEM courses, freedom from resource constraints, ease of transition for instructors) and (*b*) evidence for efficacy in promoting increased student learning (from strongest evidence, i.e., multiple high-quality comparison studies, to weakest evidence, i.e., descriptive application studies only). The following paragraphs, summarized in **Table 2**, compare these practices with their counterparts

Table 2 Comparison of traditional practices with corresponding research-based promising practices for nine aspects of large course design and implementation in STEM disciplines

Course aspect	Traditional practice	Research-based promising practice
1. Content organization	Prepare a syllabus describing the topics that the instructor will present in class.	Formulate specific student learning objectives, in the form of "after this course, students will be able to..."
2. Student organization	Most student work is done individually and competitively.	Most student work is done cooperatively in small groups.
3. Feedback	Grading based primarily or entirely on summative assessments, i.e., midterm and final exams.	Feedback to instructor and students provided continually through in-class formative assessments.
4. In-class learning activities	Instructor transmits information by lecturing. Some questions may be posed to students, but only a small subset of the class is likely to participate in discussion.	All students spend most or all class time engaged in various active-learning activities (see text) facilitated by instructors and TAs. These activities also provide formative assessment.
5. Out-of-class learning activities	Students read the text and may do assigned homework to practice application of concepts previously presented in class.	Students read and do assigned homework on new topics and post results online for the instructor to review before the class on those topics.
6. Student-faculty interaction in class	Students are expected to accept the teaching mode chosen by the instructor and to infer how they should study and what they should learn from the instructor's lectures and assignments.	Instructor explains the pedagogical reasons for the structure of course activities to encourage student buy-in, and explicitly and frequently communicates the course learning goals to students.
7. Student-faculty interaction out of class	Students must initiate out-of-class interaction with each other and with the instructor, e.g., by coming to office hours.	Instructor facilitates interaction with and among students by setting up online chat rooms, encouraging group work on homework assignments, and communicating with students electronically.
8. Use of teaching assistants (TAs)	TAs grade assignments and exams and may conduct recitation sessions to demonstrate problem solving methods or further explain lecture material.	TAs receive some initial instruction in basic pedagogy and serve as facilitators for in-class group work and tutorial sessions for small student groups to work out problems on their own.
9. Student laboratories	Students carry out exercises that demonstrate widely used techniques or verify important principles by following a prescribed protocol ("cookbook labs").	Students are required to solve a research problem, either defined (e.g., identify an unknown) or more open-ended (e.g., determine whether commonly used cosmetic products are mutagenic), and learn necessary experimental techniques and concepts in the process (inquiry-based labs).

Instructor-centered: designed around the knowledge the instructor wishes to transmit to students; focused on the instructor's teaching process

Student-centered: designed around the needs, abilities, prior knowledge, and diversity of students; focused on the student's learning process

MCAT: medical college admission test

in traditional instruction and rate them on Froyd's two criteria. The practices are organized under nine aspects of course organization.

Content Organization: The Syllabus versus Specific Learning Goals

The difference between preparing a course syllabus and formulating learning objectives is more profound than it may appear (Allen & Tanner 2007). The typical syllabus is instructor-centered; it lists the topics on which the instructor will lecture and assign out-of-class work, but it gives students little information about the level of understanding they should strive for or the skills they are to learn. In molecular biology, for example, the process of transcription can be understood at many levels, which are generally not distinguished in a syllabus. In contrast, learning objectives are student-centered and more explicit; they describe what a successful student should be able to do at the end of the course or unit. For example, students should be able to "name the principal enzyme that catalyzes transcription," "explain the nucleotide sequence relationships between the two strands of the template DNA and the RNA transcription product," "diagram a step in the elongation of an RNA transcript showing the local nucleotide sequences and strand polarities of both DNA strands and the RNA," or "predict the consequences for the transcription process if one of the four nucleoside triphosphates is unavailable."

The learning objectives above demand different levels of understanding. A half century ago, the American educator Benjamin Bloom developed a convenient scheme for classifying these levels (Bloom & Krathwohl 1956), which became known as Bloom's taxonomy of the cognitive domain (**Figure 1**). Each of Bloom's six levels of understanding can be associated with verbs appropriate for a learning goal at that level. For example, the ability to name an enzyme or describe a process requires only memorization of the relevant information (level 1), whereas ability to predict an outcome (level 3) or defend a principle based on evidence (level 6) require deeper conceptual understanding. The verbs employed (**Figure 1**) describe an action or ability that can be assessed by asking students to carry it out. Importantly, statements such as "students should understand," "appreciate," or "be aware of" are inappropriate learning objectives because their achievement cannot be tested without more explicit performance-based criteria. Because lower Bloom's levels are easier to assess with multiple-choice and short-answer exams, many instructors in large STEM courses neither demand nor test for higher levels of understanding. In a survey of over 500 final exams from a variety of introductory undergraduate and medical school biology courses, most questions were rated at Bloom's levels 1 and 2, and questions on the Medical College Admissions Test (MCAT) and Graduate Record Examination (GRE) ranked only slightly higher (Zheng et al. 2008). Another ongoing research study on assessment in introductory biology courses indicates that the overwhelming majority of test items on final exams are Bloom's level 1 (D. Ebert-May, personal communication). Because most students learn at the level assessed on summative exams, it is small wonder that they derive only superficial knowledge from such courses. Instructors can remedy this situation by aiming for higher Bloom's levels in formulating course learning goals and assessing student knowledge with appropriately challenging questions on exams (Crowe et al. 2008).

Course design around learning goals follows the principle of backward design (Wiggins & McTighe 1998). The instructor first formulates broad learning goals for students in the course and then more specific learning objectives. Once these are defined, she designs assessments (both formative and summative) to test for their achievement. Only then does she choose the most appropriate text or other reference materials and plan the learning activities in and outside of class that will most effectively lead to fulfillment of the objectives. At the start of the course, she will explicitly apprise students of the learning objectives, which may include rubrics (Allen & Tanner 2006) demonstrating

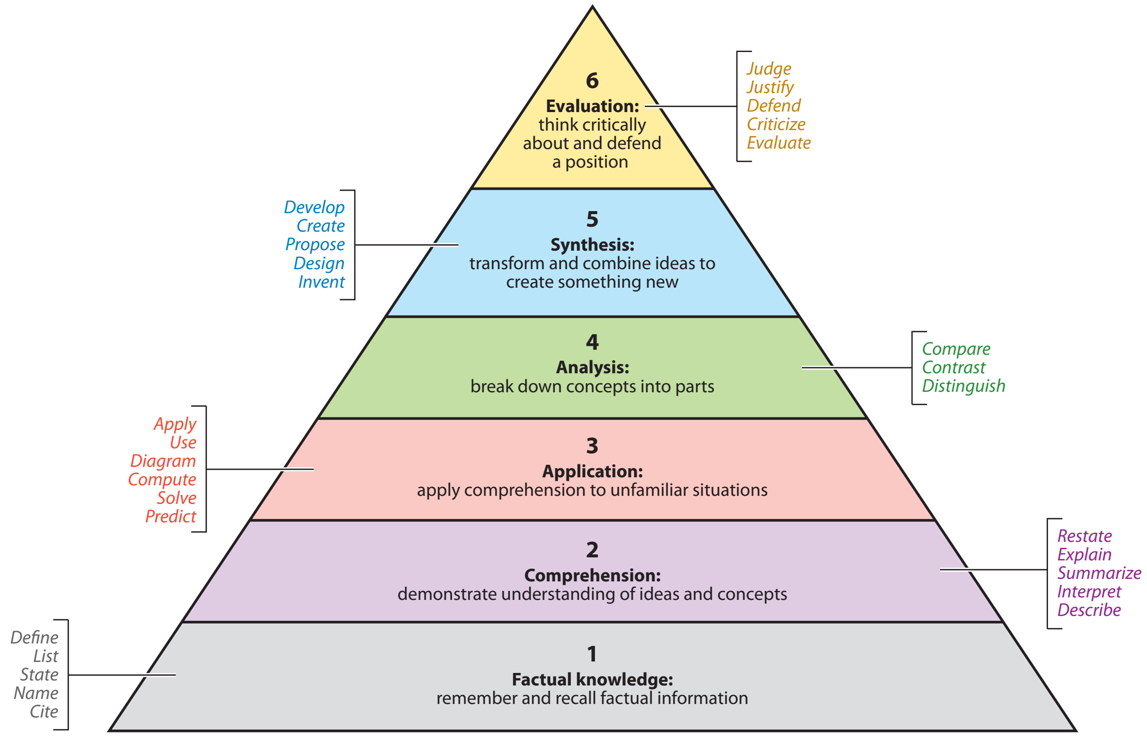


Figure 1

Bloom's levels of understanding. Originally termed Bloom's taxonomy of the cognitive domain, this schema defines six levels of conceptual understanding according to the intellectual operations that students at each level are capable of (Bloom & Krathwohl 1956). The italicized verbs have been added to the original hierarchy; they indicate performance tasks that test achievement of learning goals at each level. Fine distinctions in the hierarchy are difficult, and some educators prefer to classify goals on only three levels: low (1, 2), medium (3, 4), and high (5, 6). (Based on Allen & Tanner 2002.)

what achievement of the objectives would look like. **Figure 2** compares traditional and backward design of STEM courses.

Froyd's (2008) implementation rating for the practice of course design around learning objectives is high (applicable to any STEM course, no significant resource constraints, no need for radical change in instructor's teaching methods). As for efficacy rating, there are no empirical studies (known to this author) that compare student learning in courses taught from syllabi and those built around learning objectives. However, it seems self-evident that more learning will occur in courses that explicitly set goals for higher levels of conceptual

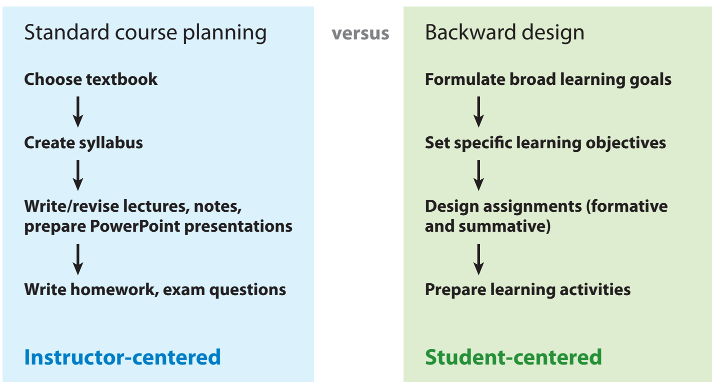


Figure 2

Schematic comparison of standard and backward course design.

understanding and require that students demonstrate achievement of these goals on exams and other course work.

Student Organization: Individual versus Group Work

Organizing students into small groups for in-class and out-of-class work can transform the course experience from competitive to collaborative, allow students to learn from each other as well as from instructors, and help to involve students who might not otherwise become actively engaged with the course content (Tanner et al. 2003). Groups can collaborate on regular homework assignments, longer-term projects such as researching a topic and developing a poster presentation, and in-class work if the course includes problem solving and other active learning activities during class time.

The implementation rating for group organization is lower than for learning objectives, because it involves additional instructor effort and decision making regarding, for example, how to form effective groups, facilitate their function, and help students develop collaborative skills (for specific references, see Froyd 2008). With regard to efficacy, much research in social science has shown that groups in general are more effective at complex problem solving than individuals (e.g., Brophy 2006) and that a group's effectiveness increases with the diversity of its members (Cox 1993, McLeod et al. 1996, Guimera et al. 2005). Comparative studies and meta-analyses provide strong evidence that group work in STEM courses contributes to increased student learning (e.g., Johnson et al. 1998, Springer et al. 1999). There is additional evidence in connection with in-class active learning in groups, discussed in the context of practice 4 below.

There are also other arguments for encouraging group work. With the increasing popularity of distance learning, the opportunity for student collaborative intellectual endeavor is one of the major advantages that resident universities can provide, and these universities should exploit it. As Astin (1993) concluded in his book of the same name, *What Matters in College* are the relationships students build with each other and with their instructors. Moreover, the development of group-work skills is important in preparing students for the real world. When students who are comfortable with the traditional individual and generally competitive learning mode object to group work, the instructor can point out that when they join the workforce, they will probably be part of a team whose members they did not choose and that they need to learn how to work effectively with a group as an important part of their education.

Feedback: Summative versus Formative Assessment

One of the key aspects of effective instruction identified in *How People Learn* (NRC 1999) is feedback to students during the learning process. Traditional courses provide feedback by returning graded homework and exams to students, often too late to be of optimal use because the class has moved on to other topics. In contrast, in-class formative assessment provides immediate feedback to both students and instructors on how well a concept under discussion is being understood. The results can be eye-opening, particularly for instructors who are considered engaging and effective lecturers, when they find that only a fraction of their students have understood a seemingly lucid explanation (see Hrepic et al. 2007). Students may be surprised as well because the concept as presented may have seemed clear until they were asked to explain or apply it. But most important, awareness of a problem in understanding allows the class to address it immediately and in context when it is most meaningful to students.

In the 1990s, the physicist Eric Mazur began to obtain this kind of feedback by posing to his class multiple-choice questions ("ConcepTests") that required application of the concept under discussion (Mazur 1997, Crouch & Mazur 2001). Initially, students indicated their choices by a show of hands or by holding up different colored cards. More recently the audience response devices known as clickers, developed originally for TV game shows, have made this kind of formative assessment more convenient and powerful (Wood 2004, Caldwell 2007, Bruff 2009). Each student has a clicker, generally with five buttons labeled A–E, and a receiver is connected to the instructor's computer. When students answer a multiple-choice question using the clickers, their answers are recorded electronically, and a histogram of the results is displayed to the instructor and, eventually, to the class. How the instructor can respond to this information is discussed in the following section on active learning, but the benefits for formative assessment are clear: Student responses are independent and anonymous, responses are recorded for later analysis by the instructor if desired, problems with understanding are immediately apparent, and the class can address these problems on the spot.

Frequent quizzes can also serve as formative assessment, and research has shown that taking tests after studying leads to significantly more learning than studying alone (Karpicke & Roediger 2008, Klionsky 2008). Moreover, the results of quizzes (and in-class concept questions) are valuable to the instructor in designing appropriate exam questions for future summative assessments. Another kind of formative assessment is the "one-minute-paper" (Angelo & Cross 1993, Stead 2005), in which students are asked to write down and hand in anonymously a brief statement of what they found most difficult and what they found most interesting during the preceding class. This exercise encourages immediate reflection on the part of students and informs the instructor of possible problems. Students can also be asked to comment, positively or negatively, about general aspects of the course. Additional types of formative assessment are considered in the following section on in-class active-engagement activities. Any activity that requires students to apply concepts just discussed can provide useful feedback about conceptual understanding to both students and instructors.

The ease of implementing formative assessment is high; instructors do not need to change the way they teach to obtain occasional feedback during class, although the results of such feedback may well change their teaching approaches as discussed further below. Clickers are an added expense for students who generally purchase a clicker at the bookstore and can resell it if they wish at the end of the course (Barber & Njus 2007). With regard to evidence for efficacy, formative assessment is generally coupled with in-class activities and so cannot be easily evaluated in isolation. Studies demonstrating the value of both these practices in combination are discussed in the following section.

In-Class Learning Activities: Listening and Note-Taking versus Active Engagement

In large STEM classes, the traditional learning activity is the lecture. Even students who are paying close attention to the lecturer are engaged primarily in the passive recording of information with little time for reflection. There is compelling evidence from all STEM disciplines that replacing some or all lecturing with in-class activities that actively engage students can substantially increase their learning gains. Of the promising practices reviewed here, this one, especially when combined with practice 2, students working in groups, and practice 3, frequent formative assessment, has produced the most impressive improvements in study after study. Many possible in-class activities—brainstorming, reflection followed by discussion with a neighbor and reporting to the class ("think-pair-share"), concept mapping, group problem solving, and more—are well described in the excellent book *Scientific Teaching* (Handelsman et al. 2007) and in the series of features titled "Approaches to Biology Teaching and Learning" by D. Allen & K. Tanner in the online journal *CBE-Life Sciences Education* (Allen & Tanner 2002; 2003a,b; 2005). **Table 3**, adapted from Handelsman et al. (2007), compares the traditional lecture presentation of a few topics with corresponding active-learning alternatives.

Table 3 Comparisons between presentation of topics in traditional lecture format and corresponding active learning activities

Concept	Passive lecture	Active class
Differential gene expression	Every cell in an organism has the same DNA, but different genes are expressed at different times and in different tissues. This is called differential gene expression.	If every cell in an animal has the same DNA, then how can cells of different tissues be so different? Discuss this question with your neighbor and generate a hypothesis.
DNA structure and replication	Complementary base pairing is the basis for the mechanism of DNA replication.	What do you know about the structure of DNA that suggests a mechanism for replication? Think about this for a minute and then discuss it with your neighbor.
Data analysis and interpretation	Based on the data shown in this slide, researchers concluded that *Snarticus inferensis* is the causal agent of the disease.	Consider these data from the experiment I just described. Which of the following conclusions can you draw from them? Think about it for a minute, and then we will take a vote and discuss the results.
Biology and society	Many people have concerns about genetically modified organisms (GMOs). Some of these concerns are well founded, and others are not. You have to decide for yourself.	I would like to split the class into two groups. One group will brainstorm about the potential benefits of GMOs and the other about possible harmful consequences. Then we will have a debate.

In-class concept questions, particularly when used with clickers, can be a powerful active learning tool. When a challenging multiple-choice concept question is presented to the class and the initial response is about evenly split between the correct choice and one or more incorrect choices (distracters), a teachable moment occurs: Students may be amused or surprised, but they want to know who is right and who is wrong, and they have become emotionally involved (Wood 2004). Rather than revealing the correct answer or trying to explain the concept again, the instructor, if interested in promoting active learning, should ask the students to discuss their answers in small groups, trying to convince their neighbors of the correct choice. Following a few minutes of discussion, the instructor calls for another vote, and almost invariably, the majority of students will now choose the correct answer, which is then revealed and discussed. Students are often better able than the instructor to identify flawed reasoning by their peers and convince them of the correct reasoning. Mazur named this phenomenon peer instruction in his delightful book of the same name (Mazur 1997, Crouch & Mazur 2001). It could be argued that less knowledgeable students are simply influenced during discussion by peer pressure from neighbors they perceive to be more knowledgeable, but a recent study indicates that, on the contrary, students are actually learning during the discussion, even when no one in a group initially knows the correct answer (Smith et al. 2009).

Distracters: the incorrect choices in a multiple-choice question

Clicker questions, to be effective, must be conceptual and challenging. Ideally they should include distracters based on known student misconceptions, and they should assess higher Bloom's levels of understanding (Modell et al. 2005, Lord & Baviskar 2007, Crowe et al. 2008). Writing good questions is challenging but essential; questions that simply test factual recall of recently presented information do not engage students and are of little pedagogical use. Clicker questions are also not helpful if the instructor, after the initial vote, simply indicates the correct answer and then moves on. Student discussion before revealing the correct answer as well as after is key to learning. For additional guidance on writing good clicker questions and their effective use, see Beatty et al. (2006), Wieman et al. (2008), and Bruff (2009).

Clicker questions generally pose well-defined, discrete problems that are directly related to the immediate class content. Other valuable problem-based activities can be based on larger, more open-ended questions that groups of students may work on for a larger fraction of the class period and continue outside of class (see following section). But all are examples of building instruction around student engagement with a problem, rather than around a body of factual information. Prince & Felder (2007) have contrasted deductive teaching—transmitting facts, abstract concepts, and finally (maybe), discussing their application to real-world problems—with inductive teaching—posing a real-world problem to students at the start, and letting them uncover the relevant concepts and facts in the process of solving it. When teaching is deductive, student motivation to learn facts and concepts is often primarily extrinsic, driven by desire to obtain a good grade, and the instructor must try to keep students engaged with assertions that this knowledge will be important in their future studies or careers. By contrast, when teaching is inductive, the students are presented with a real-world scenario (relevant to the particular group of students being taught) that they are likely to find interesting, and their motivation is intrinsic, based on desire to find a solution. Inductive approaches have been given a variety of labels including inquiry-based, problem-based, project-based, case-based, question-driven, and discovery learning (reviewed in Prince & Felder 2007). Their scope can range from a series of related clicker questions in a single class period (Beatty et al. 2006) to a complex problem requiring several weeks of work, in which new information is provided in response to requests from students for data or results of specific experimental tests. Disease-related, problem-based, and case-based activities, in which students are presented with a set of symptoms and asked to arrive at a diagnosis, are used extensively in medical education (Albanese & Mitchell 1993).

Instructors who wish to introduce more active learning into their classes may confront several problems. Implementing this mode of teaching can involve more up-front effort than the promising practices discussed above. Although designing a new course around the active learning model may require no more effort than preparing the lectures for a new traditional course, transforming a traditional course requires the additional work of creating effective in-class activities and formative assessments. Another problem is that traditional auditorium-style classrooms with fixed seating are poorly suited for interactive group work. A few institutions have installed large classrooms with café-style seating, which greatly facilitates student-centered teaching (see Beichner 2008), and more such classrooms are needed to encourage course transformation. A final problem, perhaps most difficult for some instructors, is that teaching effectively in the new mode requires both a willingness to let go of some control in the classroom and a change in perspective from instructor-centered teaching to student-centered learning. Instructors must give up the widely held transmissionist view that students must be told everything they need to know and instead realize that not only are students in a stimulating and supportive environment capable of learning a great deal on their own (the constructivist viewpoint), but that they must develop this ability in order to become either successful scientists or well-informed citizens.

Transmissionist: the view that learning can or must occur by transmission of knowledge from an instructor to the learner

Balanced against the above potential difficulties is the clear evidence from DBER that moving toward more active learning in a more student-centered classroom can substantially increase student learning gains. And complete restructuring is not necessary; even incremental changes can have a significant effect (e.g., Knight & Wood 2005). Other evidence from the life sciences has been mentioned (Udovic et al. 2002, Armstrong et al. 2007, Freeman et al. 2007), and additional references can be found in Froyd (2008).

Out-of-Class Learning Activities: Instructor versus Student Responsibility for Learning

A frequent concern of instructors contemplating introduction of clickers and other active-learning activities into their classrooms is that they will no longer be able to cover all the necessary content. First of all, this may not be a bad thing. More coverage does not necessarily mean more learning, and it can be argued that deep student understanding of a few important concepts is more valuable than superficial exposure to many concepts. Nevertheless, the content issue is real because it can affect student preparation for subsequent courses and standardized tests such as the MCAT. A solution to this dilemma lies in placing more of the responsibility on the students themselves for learning basic concepts, and again, recent technology makes this solution more practical. Using an approach that physicists have called Just-in-Time Teaching (JiTT) (Novak et al. 1999), students are assigned reading and required to submit homework online to a course Web site before a topic is considered in class. The instructor can then scan the results (sampling randomly if the class is large), determine which concepts students seem to have grasped on their own, and then focus activities in the upcoming class on concepts they found difficult. Students may resist taking this responsibility, but again, learning to do so is essential preparation for later advanced study as well as for the real world, where one cannot expect to receive a lecture whenever a new concept must be learned. An extension of JiTT, which may be more palatable to students, is the inverted classroom approach (Lage et al. 2000). Students are provided in advance of class with access to podcasts of a PowerPoint lecture by the instructor or some other multimedia presentation that serves the information transmission function of the traditional in-class lecture. Class time can then be devoted to clicker questions, solving problems, interpreting data, or other active learning activities without concerns about decreased content coverage.

The implementation of these approaches is quite simple using the Internet and one of the Web-based course management programs that are now available at most universities to instructors of large classes. Many faculty have reported not only increased student learning with these methods but also strong endorsement by students once they realized how much they were learning (e.g., Klionsky 2004, Silverthorn 2006).

In general, the practice of assigning homework is underutilized in teaching biology. Homework may not be of much help for assimilating factual information but, in transformed courses designed to help students achieve higher Bloom's levels of understanding, homework assignments that require students to practice applying concepts, solving problems, predicting outcomes, analyzing data, and designing experiments can be an invaluable supplement to similar in-class exercises. In addition to more traditional forms of homework, interactive simulations (e.g., **http://phet.colorado.edu/index.php**) and educational video games (Mayo 2009) seem likely to become increasingly useful as out-of-class learning activities.

Student-Faculty Interaction in Class: Making Pedagogy Explicit

Many students, who have become comfortable with traditional instruction, may object to the new teaching approaches and the demands that are placed on them in transformed courses: more responsibility for learning outside of class, the need to attend class regularly, the emphasis on group work, refusal of the instructor to tell them all the things they need to know, and so on. The best way to confront these objections, in the author's experience as well as in the literature (e.g., Silverthorn 2006), is to encourage buy-in by being open with students about the pedagogical reasons for new approaches and the benefits they bring. For example, the instructor can spend a few minutes introducing the concept of Bloom's levels and remind students that the skills likely to determine their success in graduate work and the job market correspond to levels 3–6, not levels 1 and 2 (**Figure 1**). Instructors can show students evidence from DBER that group work and active learning can substantially increase learning gains and point out, as mentioned above, that these activities will better prepare them for life in the real world. But instructors should also be sympathetic and supportive of students struggling with these changes, because students, like instructors, must shift their perceptions about teaching and learning in order to succeed with the new instructional approaches (Silverthorn 2006).

The active learning activities discussed above greatly increase the amount of student-faculty interaction in comparison with traditional lecture settings. Use of clickers with peer instruction, in particular, is an easy way to move classes from one-way transmission of information to interactive dialogs between instructor and students, and between students, with instructional benefits that have been documented by research as described in the preceding paragraphs.

Student-Faculty Contact Out of Class: Office Hours versus Enhanced Communication

Umbach & Wawrzynski (2005) cite several studies showing that, in general, student learning is enhanced by increased student-faculty contact, suggesting that faculty, as time permits, should provide more opportunities for interaction than simply holding office hours for those (often few) students who will make use of them. Additional interactions can include brief get-acquainted visits by invitation to the instructor's office or, for larger courses, virtual communication through emails to the class, moderated discussion forums, or use of social networking Web sites. Aside from requiring some additional faculty time, this practice is easy to implement and its efficacy is supported by the studies referenced above.

Use of Teaching Assistants: Grading and Recitation versus Facilitation of Student Learning

Many instructors of large STEM courses have help from one or more teaching assistants (TAs), whose principal tasks are grading of homework and exams and perhaps conducting recitation sessions to go over lecture material and solutions to homework problems. If TAs are made part of the course transformation process and given minimal pedagogical training (e.g., reading of Handelsman et al. 2007), they can serve as valuable facilitators in class for discussion of clicker questions or group work on problems. In addition, they will have gained a new kind of teaching experience that can serve them well in the future if they should go on to become faculty themselves. Many institutions, for example those involved in the Center for the Integration of Research, Teaching, and Learning (CIRTL) Network (**http://www.cirtl.net/**), provide such training to STEM graduate students in Preparing Future Faculty programs (e.g., Miller et al. 2008). This practice is quite easy to implement, and research to evaluate its efficacy is in progress at the author's institution and elsewhere (personal communications).

Student Laboratories: Cookbook Exercises versus Inquiry

As one solution to the problem of inadequate STEM education for undergraduates, the Carnegie Foundation's Boyer Commission Report (Boyer 1998) recommended that research universities integrate their research and teaching missions by involving more students in the process of research. In traditional "cookbook labs" associated with many large introductory lecture courses, students perform prescribed exercises in which they may learn some laboratory techniques but generally gain little understanding of scientific inquiry. At the other end of the lab experience spectrum (see **Figure 3**), some undergraduates become apprentices in faculty laboratories, learning how science is done by working alongside graduate students and postdocs on research projects that often result in publication. Although this experience is highly desirable, most departments can provide it to only a fraction of their majors. Between these extremes, some departments have developed a variety of inquiry-based laboratory courses designed to introduce large numbers of students to the process of research (reviewed in Weaver et al. 2008). These courses range from guided inquiry labs to open-ended group research projects that may result in publications by undergraduates (e.g., Hanauer et al. 2006). Faculty who supervise these courses often design them to yield results that can contribute directly to their own research programs.

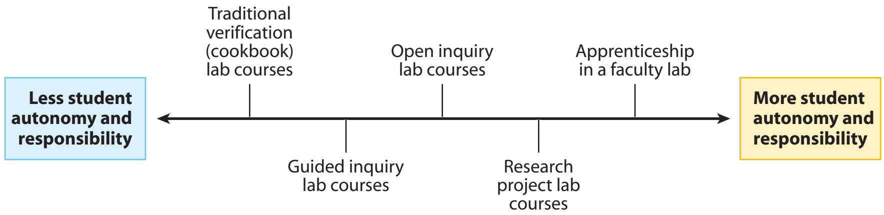


Figure 3

The range of student laboratory experiences from verification exercises ("cookbook labs") to apprenticeship in a faculty research laboratory. Levels of student responsibility and autonomy increase from left to right. (Adapted from Weaver et al. 2008.)

Implementation of inquiry-based courses in place of traditional labs may require additional resources including more extensive training for TAs. Although Froyd (2008) rates this promising practice low in terms of evidence for efficacy, several studies, in addition to the two cited above, have shown that engagement of students with real research problems is one of the most effective ways to move students along the path from novice to expert (Nagda et al. 1998, Lopatto 2004, Luckie et al. 2004, Seymour et al. 2004). Compared to students who experience only traditional lab courses, reported benefits to students in inquiry-based curricula include deeper understanding of content, increased confidence in their ability to understand and perform science, more positive attitudes about science, and lower attrition rates. These gains are particularly evident among underrepresented minority students (Nagda et al. 1998, Russell et al. 2007). Thus, the benefits of this promising practice can include not only increased student learning and higher retention of students in the major (especially if inquiry-based labs are introduced early in the curriculum) but also contributions to faculty research.

CONCLUSION: THE DUAL FUNCTIONS OF BIOLOGY EDUCATION

There are two important purposes for the introductory biology courses we teach. One is to attract, motivate, and begin preparing the next generation of biologists including the research stars of the future. The other is to help the large majority of our students who will not become biologists or even scientists to achieve minimum biological literacy and to understand the nature of science, the importance of empirical evidence, and the basic principles that underlie biological systems. They will need this knowledge as twenty-first century citizens of the United States and the world to make intelligent decisions about issues such as personal health, conflicting claims in the media, energy policy, climate change, and conservation of natural resources.

Traditional teaching methods do not prevent the progress of superior students from introductory courses to upper-level courses to graduate training, where they may become experts in their fields and develop into skilled researchers. But the traditional methods fail the majority of students who leave our introductory courses viewing biology as a large collection of disconnected facts that have little relevance to their daily lives and will soon be forgotten. Part of the problem, as described in this review, lies not in what we teach these students (though this is also a concern; see NRC 2003, AAMC-HHMI 2009) but in how we teach it. We must do better! Widespread adoption of the research-based promising practices described here will help.

SUMMARY POINTS

1. We must improve the undergraduate teaching of biology and other STEM disciplines to remain competitive in the global economy and educate American citizens adequately.
2. Recent research in educational psychology, cognitive science, and neurobiology has yielded important new insights into how people learn and the optimal conditions for learning.
3. Discipline-based educational research (DBER) has led to the development of teaching approaches based on these insights (promising practices) and has provided extensive evidence that these approaches can be substantially more effective than traditional lecturing even in large classes.
4. These promising practices vary in their ease of implementation but even their partial adoption can lead to significant gains in student learning.
5. Applying these promising practices widely in STEM classes can have a major impact on better preparing our undergraduate biology students for their future endeavors.

DISCLOSURE STATEMENT

The author is Editor-in-Chief of *CBE-Life Sciences Education*.

LITERATURE CITED

AAMC-HHMI Committee. 2009. *Scientific Foundations for Future Physicians*. Washington, DC: Assoc. Amer. Med. Coll. and Howard Hughes Med. Inst.

Albanese M, Mitchell S. 1993. Problem-based learning: a review of the literature on its outcomes and implementation issues. *Acad. Med.* 68:52–81

Allen D, Tanner K. 2002. Approaches to cell biology teaching: questions about questions. *Cell Biol. Educ.* 1:63–67

Allen D, Tanner K. 2003a. Approaches to cell biology teaching: learning content in context-problem–based learning. *Cell Biol. Educ.* 2:73–81

Allen D, Tanner K. 2003b. Approaches to cell biology teaching: mapping the journey-concept maps as signposts of developing knowledge structures. *Cell Biol. Educ.* 2:133–36

Allen D, Tanner K. 2005. Infusing active learning into the large-enrollment biology class: seven strategies, from the simple to complex. *Cell Biol. Educ.* 4:262–68

Allen D, Tanner K. 2006. Rubrics: tools for making learning goals and evaluation criteria explicit for both teachers and learners. *CBE-Life Sci. Educ.* 5:197–203

Allen D, Tanner K. 2007. Putting the horse back in front of the cart: using visions and decisions about high-quality learning experiences to drive course design. *CBE-Life Sci. Educ.* 6:85–89

Anderson DL, Fisher KM, Norman GJ. 2002. Development and evaluation of the conceptual inventory of natural selection. *J. Res. Sci. Teach.* 39:952–78

Angelo TA, Cross PK. 1993. *Classroom Assessment Techniques: A Handbook for College Teachers*. San Francisco: Jossey-Bass

Armstrong N, Chang SM, Brickman M. 2007. Cooperative learning in industrial-sized biology classes. *CBE-Life Sci. Educ.* 6:163–71

Astin AW. 1993. *What Matters in College: Four Critical Years Revisited*. San Francisco: Jossey-Bass

Ausubel D. 1963. *The Psychology of Meaningful Verbal Learning*. New York: Grune and Stratton

Ausubel D. 2000. *The Acquisition and Retention of Knowledge: A Cognitive View*. Boston: Kluwer Acad. Publ.

Barber M, Njus D. 2007. Clicker evolution: seeking intelligent design. *CBE-Life Sci. Educ.* 6:1–8

The articles by Allen & Tanner, plus another by Tanner et al. 2003, provide excellent practical introductions to several of the promising practices discussed in the text.

Beatty ID, Leonard WJ, Gerace WJ, Dufresne RJ. 2006. Question driven instruction: teaching science (well) with an audience response system. In *Audience Response Systems in Higher Education: Applications and Cases*, ed. DA Banks. Hershey, PA: Inf. Sci. Publ.

Beichner R. 2008. *The SCALE-UP Project: a student-centered active learning environment for undergraduate programs*. Presented at BOSE Conf. Promis. Pract.—Innov. Undergrad. STEM Educ., Washington, DC. **http://www7.nationalacademies.org/bose/PP_Commissioned_Papers.html**

Bloom BS, Krathwohl DR. 1956. *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 1: Cognitive Domain*. New York: Longmans

Bowling BV, Acra EE, Wang L, Myers MF, Dean GE, et al. 2008. Development and evaluation of a genetics literacy assessment instrument for undergraduates. *Genetics* 178:15–22

Boyer Commission on educating undergraduates in the research university. 1998. *Reinventing Undergraduate Education: A Blueprint for America's Research Universities*. Stony Brook, NY: Carnegie Found.

Brophy DR. 2006. A comparison of individual and group efforts to creatively solve contrasting types of problems. *Creativity Res. J.* 18:293–315

Bruff D. 2009. *Teaching with Classroom Response Systems: Creative Active Learning Environments*. San Francisco: Jossey-Bass

Caldwell JE. 2007. Clickers in the large classroom: current research and best-practice tips. *CBE-Life Sci. Educ.* 6:9–20

Cox J. 1993. *Cultural Diversity in Organizations: Theory, Research, and Practice*. San Francisco: Berrett-Koehler Publ.

Crouch CH, Mazur E. 2001. Peer instruction: ten years of experience and results. *Am. J. Phys.* 69:970–77

Crowe A, Dirks C, Wenderoth MP. 2008. Biology in bloom: implementing Bloom's taxonomy to enhance student learning in biology. *CBE-Life Sci. Educ.* 7:368–81

DeHaan RL. 2005. The impending revolution in undergraduate science education. *J. Sci. Educ. Technol.* 14:253–59

Dewey J. 1916. *Democracy and Education*. New York: Macmillan

DoE. 2000. Before it's too late (the Glenn Report). Dept. Educ., Washington DC. **http://www.ed.gov/americacounts/glenn**

Dolan EL. 2007. Grappling with the literature of education research and practice. *CBE-Life Sci. Educ.* 6:289-96

Freeman S, O'Connor E, Parks JW, Cunningham M, Hurley D, et al. 2007. Prescribed active learning increases performance in introductory biology. *CBE-Life Sci. Educ.* 6:132–39

Froyd J. 2008. *White paper on promising practices in undergraduate STEM education*. Presented at BOSE Conf. Promis. Pract.—Innov. Undergrad. STEM Educ., Washington, DC. http://www7.nationalacademies.org/bose/PP_Commissioned_Papers.html

A good summary of the evidence with additional references for most of the promising practices described in this review.

Garvin-Doxas K, Klymkowsky M, Elrod S. 2007. Building, using, and maximizing the impact of concept inventories in the biological sciences: report on a National Science Foundation sponsored conference on the construction of concept inventories in the biological sciences. *CBE-Life Sci. Educ.* 6:277–82

Guimera R, Uzzi B, Spiro J, Amaral LA. 2005. Team assembly mechanisms determine collaboration network structure and team performance. *Science* 308:697–702

Hake RR. 1998. Interactive-engagement versus traditional methods: a six-thousand-student survey of mechanics test data for introductory physics courses. *Am. J. Phys.* 66:64–74

Halloun IA, Hestenes D. 1985. Common-sense concepts about motion. *Am. J. Phys.* 53:1056–65

Hanauer DI, Jacobs-Sera D, Pedulla ML, Cresawn SG, Hendrix RW, Hatfull GF. 2006. Inquiry learning. Teaching scientific inquiry. *Science* 314:1880–81

Handelsman J, Ebert-May D, Beichner R, Bruns P, Chang A, et al. 2004. Scientific teaching. *Science* 304:521–22

Handelsman J, Miller S, Pfund C. 2007. *Scientific Teaching*. New York: W.H. Freeman

An excellent, brief practical guide to most of the promising practices described in this review. Highly recommended for all STEM instructors.

Hestenes D, Wells M, Swackhamer G. 1992. Force concept inventory. *Phys. Teach.* 30:141–58

Holt J. 1964. *How Children Fail*. New York: Pitman

Holt J. 1967. *How Children Learn*. New York: Pitman

Hrepic Z, Zollman DA, Rebello SJ. 2007. Comparing students' and experts' understanding of the content of a lecture. *J. Sci. Educ. Technol.* 16:213–24

Johnson DW, Johnson RT, Smith KA. 1998. Cooperative learning returns to college: What evidence is there that it works? *Change* 30:26–35

Karpicke JD, Roediger HL III. 2008. The critical importance of retrieval for learning. *Science* 319:966–68

Klionsky D. 2004. Talking biology: teaching outside the textbook, and the lecture. *Cell Biol. Educ.* 3:204–11

Klionsky D. 2008. The quiz factor. *CBE-Life Sci. Educ.* 7:265–66

Klymkowsky MW, Garvin-Doxas K, Zeilik M. 2003. Bioliteracy and teaching efficacy: what biologists can learn from physicists. *Cell Biol. Educ.* 2:155–61

Knight JK, Wood WB. 2005. Teaching more by lecturing less. *Cell Biol. Educ.* 4:298–310

Kozol J. 1967. *Death at An Early Age*. Boston: Houghton-Mifflin

Lage MJ, Platt GJ, Treglia M. 2000. Inverting the classroom: a gateway to creating an inclusive learning environment. *J. Econ. Educ.* 31:30–43. **http://www.sba.muohio.edu/plattgj/eco201**

Libarkin J. 2008. Concept inventories in higher education science. Presented at BOSE Conf. Promis. Pract.—Innov. Undergrad. STEM Educ., Washington, DC. **http://www7.nationalacademies.org/bose/PP_Commissioned_Papers.html**

Lopatto D. 2004. Survey of undergraduate research experiences (SURE): first findings. *Cell Biol. Educ.* 3:270–77

Lord T, Baviskar S. 2007. Moving students from information recitation to information understanding: exploiting Bloom's taxonomy. *J. Coll. Sci. Teach.* 36:40–44

Luckie DB, Maleszewski JJ, Loznak SD, Krha M. 2004. Infusion of collaborative inquiry throughout a biology curriculum increases student learning: a four-year study of "Teams and Streams." *Adv. Physiol. Educ.* 28:199–209

Mayo MJ. 2009. Video games: a route to large-scale STEM education? *Science* 323:79–82

Mazur E. 1997. *Peer Instruction: A User's Manual*. Upper Saddle River, NJ: Prentice Hall

McLeod PO, Lobel SA, Cox TH. 1996. Ethnic diversity and creativity in small groups. *Small Group Res.* 27:248–65

Miller S, Pfund C, Pribbenow CM, Handelsman J. 2008. Scientific teaching in practice. *Science* 322:1329–30

Modell HM, Michael J, Wenderoth MP. 2005. Helping the learner to learn: the role of uncovering misconceptions. *Am. Biol. Teach.* 67:20–26

Nagda BA, Gregerman SR, Jonides J, von Hippel W, Lerner JS. 1998. Undergraduate student-faculty research partnerships affect student retention. *Rev. Higher Educ.* 22:55–72

Natl. Acad. Sci. 2004. *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*. Washington DC: Natl. Acad. Press. **http://books.nap.edu/catalog/11463.html**

Natl. Acad. Sci. 2005. Natl. Acad. Worksh. Educ. Res. Positions STEM Discipl. Dep. Washington, DC. **http://www7.nationalacademies.org/cfe/STEM_Disciplines_Agenda.html**

Natl. Res. Counc. 1999. *How People Learn: Brain, Mind, Experience, and School.* Bransford JD, Brown AL, Cocking RR, eds. Washington, DC: Natl. Acad. Press

Classic review of educational and cognitive science research on how learning occurs and the conditions that foster it.

Natl Res. Counc. 2003. *Bio2010: Transforming Undergraduate Education for Future Research Biologists*. Stryer L, ed. Washington, DC: Natl. Acad. Press

Novak G, Gavrin A, Christian W, Patterson E. 1999. *Just-in-Time Teaching: Blending Active Learning with Web Technology*. San Francisco: Benjamin Cummings. **http://webphysics.iupu.edu/jitt/jitt.html**

Prince M, Felder R. 2007. The many faces of inductive teaching and learning. *J. Coll. Sci. Teach.* 36:14–20

Review of instructional approaches that present real-world problems to students and challenge them to discover the facts and concepts they need for a solution.

Russell SH, Hancock MP, McCullough J. 2007. The pipeline. Benefits of undergraduate research experiences. *Science* 316:548–49

Seymour E, Hewitt S. 1997. *Talking About Leaving: Why Undergraduates Leave the Sciences*. Boulder, CO: Westview Press

Seymour E, Hunter AB, Laursen SL, Diatonic T. 2004. Establishing the benefits of research experiences for undergraduates in the sciences: first findings from a three-year study. *Science Educ.* 88:493–534

Silverthorn DU. 2006. Teaching and learning in the interactive classroom. *Adv. Physiol. Educ.* 30:135–40

Smith AC, Stewart R, Shields P, Hayes-Klosteridis J, Robinson P, Yuan R. 2005. Introductory biology courses: a framework to support active learning in large enrollment introductory science courses. *Cell Biol. Educ.* 4:143–56

Smith MK, Wood WB, Adams WK, Wieman C, Knight JK, et al. 2009. Why peer discussion improves student performance on in-class concept questions. *Science* 323:122–24

Smith MK, Wood WB, Knight JK. 2008. The genetics concept assessment, a new concept inventory for gauging student understanding of genetics. *CBE-Life Sci. Educ.* 7:422–30

Springer L, Stanne ME, Donovan SS. 1999. Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis. *Rev. Educ. Res.* 69:21–51

Stead DR. 2005. A review of the one-minute paper. *Active Learn. Higher Educ.* 6:118–31

Tanner K, Chatman LS, Allen D. 2003. Approaches to cell biology teaching: cooperative learning in the science classroom—beyond students working in groups. *Cell Biol. Educ.* 2:1–5

Udovic D, Morris D, Dickman A, Postlethwait J, Wetherwax P. 2002. Workshop biology: demonstrating the effectiveness of active learning in an introductory biology course. *Bioscience* 52:272–81

Umbach PD, Wawrzynski MR. 2005. Faculty do matter: the role of college faculty in student learning and engagement. *Res. Higher Educ.* 46:153–84

Weaver GC, Russell DB, Wink DJ. 2008. Inquiry-based and research-based laboratory pedagogies in undergraduate science. *Nature Chem. Biol.* 4:577–80

Wieman C. 2007. Why not try a scientific approach to science education? *Change* 39:9–15

Wieman C, Perkins K, Gilbert S, Benay F, Kennedy S, et al. 2008. *Clicker Resource Guide: An Instructors Guide to the Effective Use of Personal Response Systems (Clickers) in Teaching*. Vancouver, BC: Univ. British Columbia. http://www.cwsei.ubc.ca/resources/files/Clicker_guide_CWSEI_CU-SEI_04-08.pdf

Wiggins G, McTighe J. 1998. *Understanding by Design*. Alexandria, VA: Assoc. Superv. Curric. Dev.

Wood WB. 2004. Clickers: a teaching gimmick that works (resource review). *Dev. Cell* 7:796–98

Zheng AY, Lawhorn JK, Lumley T, Freeman S. 2008. Application of Bloom's taxonomy debunks the "MCAT myth." *Science* 319:414–15

Zull JE. 2002. *The Art of Changing the Brain*. Sterling, VA: Stylus

A good review of inquiry-based curricular innovations and their utility for introducing more students to the process of research.

A useful introductory guide to clickers and their effective use to promote active learning through in-class concept questions.

A book for educators on the neurobiology of learning.