Taking a Scientific Approach to Science Education, Part II—Changing Teaching

Challenges remain before universities more widely adopt research-based approaches, despite their many benefits over lecture-based teaching

Carl Wieman and Sarah Gilbert

In Part I of "Taking a Scientific Approach to Science Education" (*Microbe*, April 2015, p. 152) we focused on the research on learning and teaching in undergraduate science. Our focus here is to address why these improved teaching methods are not the norm in college and university science classes, and what it will take to achieve widespread adoption.

In Part I we discussed the cognitive psychology research on the nature of expertise and how it is acquired. We then mapped those research findings onto the learning and teaching of science, and presented examples of research in university science courses. In those examples, learning outcomes were compared between classes using conventional lecture methods and those that implemented the "practice of expert thinking with feedback" that cognitive psychology has shown is essential for developing expertise. These examples illustrate the large increases in learning that result with the research-based teaching methods and substantial reductions in failure rates.

In addition to examples described in Part I, there is a vast literature of similar studies across the science and engineering disciplines. These results are summarized in a meta-analysis of 225 papers by Scott Freeman and coauthors (Fig.1). They compared failure rates and performance on identical or near-identical exams for courses that were taught using traditional lecture methods or incorporating "active learning" methods (our "authentic practice and feedback"). On average, in active learning courses, the failure rates are 35% lower, the exam scores are 0.47 standard deviations higher, and scores on concept inventory tests are 0.9 standard deviations higher. The larger difference on the latter type of test is to be expected, as concept inventories are carefully developed research instruments that specifically target the extent to which students learned to think like experts in the discipline.

As discussed in Part I, active learning with feedback has been found to be necessary for developing expert thinking. According to Freeman and his collaborators, these benefits of active learning methods over lecture instruction are consistent across all fields of science and engineering and all course levels and study designs. They also suggest that "more is better," in that those studies with results that were well above the average also relied more extensively on active learning methods.

Systemic Change in Teaching Methods

Moving from research demonstrations to routine widespread use of these more effective teaching methods remains challenging. To understand why, one must consider the general culture of

SUMMARY

- The benefits of active learning methods over lecture instruction are consistent across all fields of science and engineering and all course levels.
- In scaling up demonstrations of active-learning experiments, it is essential to work at the department level, since departments determine what and how topics in their discipline are taught.
- Hiring teaching and learning fellows addresses the first challenge to implementing new teaching methods: members of the faculty typically lack expertise in those methods and time to learn about them.
- A critical step needed to make teaching effectiveness a more prominent component of the university incentive system is to have adequate means for gauging the quality of teaching.
- Several simple tools are helping to make the evaluation of science teaching more rigorous and meaningful.

higher education and what structural factors need to be addressed to bring about change.

The current system of teaching by lecture started thousands of years ago, when it was the only means for transmitting information from one to many. We argue that this function became obsolete with the invention of the printing press, but traditions are slow to change. Two factors likely play a part in the continuation of the lecture for teaching science. First, "teaching by telling" does work under some circumstances: when the learning is simple and obvious or the brain is well prepared to receive and process that knowledge. An example of the latter is experts learning of the latest research developments in their area of expertise. However, it is only through well-designed research studies that it is possible to determine when and where this approach fails, as it does when attempting to teach science expertise to relatively novice brains.

The second reason is that the human brain has evolved to learn, and so learning can happen largely independent of instruction, and a few success stories may appear to validate the teaching method. This situation is similar to medicine, where for 2,000 years bloodletting was considered to be an effective treatment for illness because many people that received such a treatment recovered—not because the treatment was effective, but because we have a highly evolved and effective immune system.

In the modern world, the original function of the classroom, simple information transfer, has been replaced by books and the Internet. The classroom has now become the best opportunity for students to practice the desired thinking while getting timely, specific feedback needed to support learning. This feedback comes from interactions with their fellow students and teachers. While technology can support this learning process, current technology does not yet provide the extensive engagement, interactions, and timely targeted feedback provided in a well-run classroom.

Scaling up Starts at the Department Level

The first step in scaling up these demonstration experiments must be department-level change. Departments are the units that determine what will be taught in their discipline at universities and how that will be taught. We led an experiment at a large research-intensive public university aimed at achieving department-wide changes in teaching practices. This effort had substantial activities in six departments, including biology. Although we are in the early process of writing up an extensive discussion of this experiment, here we offer a few relevant findings.

It is possible to achieve large-scale change in teaching. Our program led to substantial changes in the teaching of more than 160 courses providing about 150,000 credit hours per year to 15,000 students. This is more than half the credit hours provided by the College of Science and up to 75% of the annual credits provided by the departments that our program targeted. In two departments, nearly 90% of the faculty changed the way they teach.

This experience showed us that with the right support and motivation, nearly all faculty can learn to use more effective teaching methods. It does take time to change. Roughly 100 hours of practice is needed to switch to using new teaching methods effectively. We see that once faculty members have switched teaching approaches in one course, they can—and typically do—readily adopt these methods for other courses that they teach.

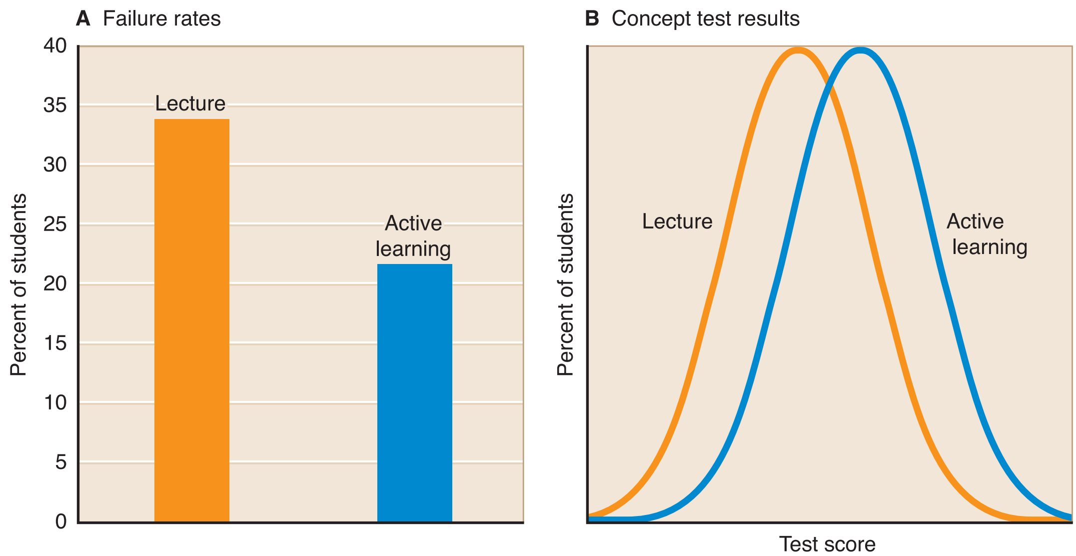
Factors That Help or Hinder Change

Our program was largely an experiment in organizational change, rather than a model intended to be replicated at other institutions. Here we will focus on the factors common to many universities that can help or hinder teaching improvement.

We give only a very brief overview of our own program to set the context. With university and donor funds, we set up a program that provided substantial one-time grants of as much as $1.75 million over six years on a competitive basis to each selected department to change how their courses are taught. Each department had considerable flexibility in how it used its grant money. However, they all chose to use the bulk of that money to hire short-term science teaching and learning fellows (STLFs)—typically, new postdocs in the discipline who then received training in science education.

Hiring such fellows addresses the first challenge to implementing new teaching methods: members of the faculty typically lack expertise in those methods and time to learn about them. The STLFs provided expertise, assistance, and coach-

FIGURE 1



Freeman et al. meta-analysis results. (From C. Wieman, Proc. Natl. Acad. Sci. USA. 111: 8319–8320, 2014.)

ing while the faculty member was first designing and teaching a transformed course, thereby minimizing the time needed for faculty to learn new methods and develop new materials for the course. The fellows also greatly reduced—in practice, eliminated—the risk of failure for the faculty member.

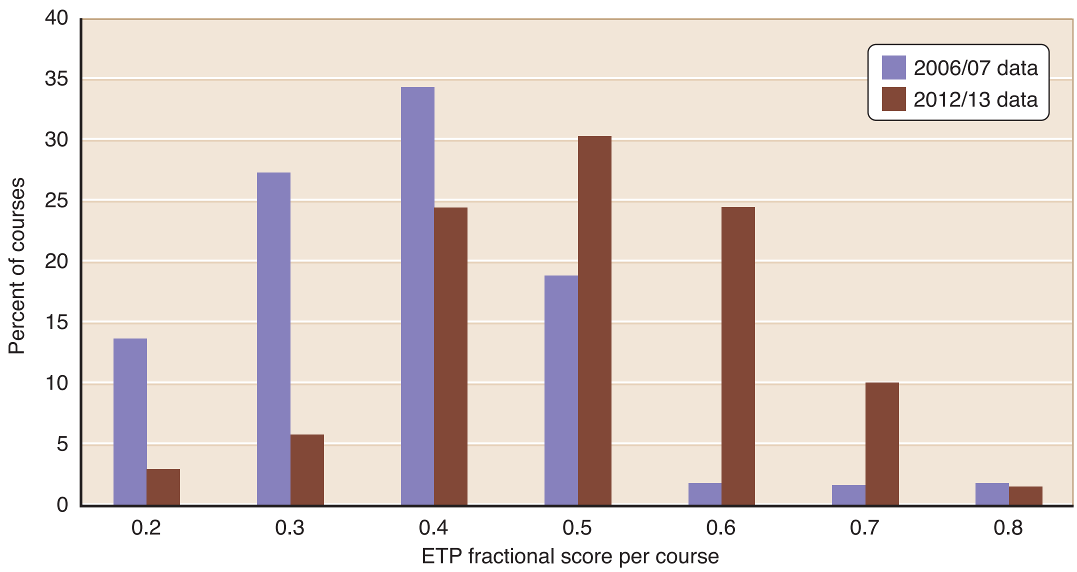
Several other things also proved important for implementing these teaching changes:

- Although data will get the attention of faculty members, they tend not to find such information convincing unless it is about their students in their classes. Because few faculty members believe themselves to be poor teachers, they search for reasons to explain that research studies of other courses or institutions do not apply to them.
- What many faculty members find most convincing is to see a class in their department taught with research-based methods, with much more engaged students who are keenly interested in the subject and in learning more about it.
- The degree to which teaching and undergraduate student outcomes are made an explicit significant part of the regular departmental function, particularly by the department chair, has a significant effect on acceptance of these new methods. Supportive department chairs regularly brought up teaching at faculty meetings, and charged committed and effective people with the responsibility to oversee improvements in teaching.
- The incentive system is the dominant factor in determining the attention faculty members pay to undergraduate teaching. This system penalizes faculty for taking time away from research to learn and implement new teaching methods, and it ignores what teaching methods faculty members use.

Toward Changing the Incentive System: Meaningful Evaluation of Science Teaching

While lip service is paid to the importance of good teaching, in an extensive survey we have found no university that collects data on what teaching methods are used in their classrooms. The primary method for evaluating teaching is student course evaluations, which are extremely limited in the information they provide, and, if anything, discourage faculty from adopting better teaching practices. Further, these evaluations correlate more strongly with factors—such as class size, level, and elective versus required course—that lie outside an instructor's control than with any measures of learning. Also, student evaluations provide little or no guidance for how to improve teaching.

FIGURE 2



Histogram of the fractional ETP scores (a gauge of the extent of use of research-based teaching practices) in the 2006–2007 and 2012–2013 academic years for the courses in a science department. The department participated in the science education initiative beginning in 2007. (Adapted from C. Wieman and S. Gilbert, CBE-Life Sci. Edu. 13:552–569, 2014.)

A crucial step needed to make teaching effectiveness a more prominent component of the university incentive system is to have adequate means for gauging the quality of teaching. As part of our initiative, we developed two tools to allow meaningful measures of teaching. The primary tool is the Teaching Practices Inventory, which allows individual instructors or departments to easily characterize their teaching approaches. This survey takes an individual faculty member about 10 minutes per course to complete and characterizes the full range of methods used in a course.

This inventory provides a simple way for individual faculty and departments to measure the quality of their teaching, find opportunities for improving it, and document that improvement. The inventory consists of eight categories covering the full range of decisions and activities that go into teaching nearly any science (and most other) courses. They include: course information, supporting materials, in-class features and activities, assignments, feedback and testing, other items such as use of instructor-independent tests to measure learning, training and guidance of teaching assistants, and collaboration or sharing in teaching. We have also developed a rubric that translates this information into a measure of the extent of use of research-based methods in a particular course.

For example, the inventory helped to document the results of a successful change in teaching practices carried out by one large science department, by quantifying the extent of their use of effective teaching practices (the ETP score in Fig. 2). The improvement over the six-year period is very clear. The range across individual courses in this department is large and typical of all the departments that we measured.

The second tool is the Classroom Observation Protocol for Undergraduate STEM (COPUS). This protocol allows individuals with 1.5 hours of training to make reliable and consistent observations as to how an instructor and students are spending their time in a classroom. An increasing number of universities are using the COPUS and/or the Teaching Practices Inventory to examine the teaching of a subset of their courses.

As such practices make evaluation of science teaching more rigorous and meaningful, we are optimistic that this trend will drive more universities and colleges to adopt research-based teaching methods. This result would lead to science teaching becoming more a science itself, like astronomy, rather than being determined by habit and superstition, like astrology. The benefits to students and society of such a change will be enormous.

Carl Wieman holds a joint appointment as Professor of Physics and of the Graduate School of Education at Stanford University, Stanford, Calif., and Sarah Gilbert is a senior advisor at the Carl Wieman Science Education Initiative, University of British Columbia, Vancouver, Canada. Part 1 of this feature appeared in the April 2015 issue of Microbe.

Suggested Reading

Carl Wieman Science Education Initiative. 2015. Videos showing research-based methods: http://www.cwsei.ubc.ca/resources/SEI_video.html.

Freeman, S., S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth. 2014. Active learning increases student performance in science, engineering, and mathematics. Proc. Natl. Acad. Sci. USA **111:**8410 – 8415.

iBiology.org. 2014. Scientific teaching video series—active learning module: http://www.ibiology.org/scientific-teaching/active-learning.html.

National Research Council. 2012. Discipline-based education research: understanding and improving learning in undergraduate science and engineering. National Academies Press, Washington, D.C.

Smith, M., F. Jones, S. Gilbert, and C. Wieman. 2013. The classroom observation protocol for undergraduate STEM (COPUS): a new instrument to characterize university STEM classroom practices. CBE-Life Sci. Edu. **12:**618 – 627.

Smith, M., E. Vinson, J. Smith, J. Lewin, and M. Stetzer. 2014. A campus-wide study of STEM courses: new perspectives on teaching practices and perceptions. CBE-Life Sci. Edu. **13:**624 – 635.

Wieman, C., K. Perkins, and S. Gilbert. 2010. Transforming science education at large research universities: a case study in progress. Change **42**(2)**:**7–14.

Wieman, C., and S. Gilbert. 2014. The teaching practices inventory: a new tool for characterizing college and university teaching in mathematics and science. CBE-Life Sci. Edu. **13:**552–569.

Wieman, C. 2015. A better way to evaluate undergraduate teaching. Change **47**(1)**:**6 –15.