

Why the Science of Teaching Is Often Ignored

There's a whole literature on what works. But it's not making its way into the classroom.

The Chronicle of Higher Education

Beth McMurtrie

Jan 3, 2022

A couple of years ago, five faculty members at Harvard University published an intriguing study. They had run an experiment in an introductory undergraduate physics course to figure out why active learning, a form of teaching that has had measurable success, often dies a slow death in the classroom. (Deslauriers, et al. (2019). "Measuring Actual Learning Versus Feeling of Learning in Response to Being Actively Engaged in the Classroom." Proceedings of the National Academy of Sciences 116(39): 19251.)

The authors compared the effects of a traditional lecture with the effects of active learning, in which students solve problems in small groups. They found — to little surprise — that when students were taught in an active format they performed better on tests. Then they made another, more striking, discovery: Students felt like they were learning more when they sat through a lecture. In other words, though they were very engaged by the talk, it didn't actually help them understand physics better.

Academic Twitter praised the study for its clever design and for the way it resonated with professors who had struggled with active learning. But even as it was lauded in some quarters, the study was picked apart in others. It measured the effects of single lessons, some complained. Could you really conclude, others asked, that one test was a true measure of learning? The experiment said nothing about long-term retention, still other critics pointed out. Would those differences in scores still be apparent months later?

That mixed reaction illuminates a central paradox in higher education. Scholarship on teaching and learning has grown exponentially over the decades, encompassing thousands of experiments, stacks of books and journal articles, and major initiatives to bring the science of learning into the classroom. Yet many faculty members are untouched by this work, unsure how to apply it to their teaching, or skeptical of its value.

To be sure, many instructors have participated in workshops run by their campus teaching centers. And the use of some evidence-based teaching practices, such as peer learning or the use of clickers to keep students engaged in the classroom, are far more prevalent than they were a generation ago. But faculty developers, education researchers, and learning scientists say they often feel like they are speaking to a select audience: namely, each other, or the same subset of professors eager to try new practices. And what does get through to many faculty members and students is often garbled, or just one piece of the puzzle.

So what's going on? Some of the bottlenecks are a product of the structures and systems of higher education, in which faculty members are given few incentives for, if not actively discouraged from, improving their teaching. They care about their students, but they don't have the time, understanding, or motivation to make their courses better. And if habits and preconceived notions about teaching remain unchallenged, say teaching experts, there's little reason to change.

But it's more complicated than that. Much of the research on teaching and learning is done on a small scale, perhaps in a single classroom or a lab-based experiment. How it might apply in different contexts, with different groups of students, isn't always clear. Does the success of group work in an introductory physics class, for instance, say anything about how to run a Shakespeare seminar? Students, after all, are not interchangeable variables and classrooms are not laboratories.

This confusion and discomfort are also partly a natural consequence of the relative youth of the field. It's messy and not very definitive. Classroom experiments may be flawed.

Yet, teaching reformers argue, the dangers of ignoring the expanding body of knowledge about teaching and learning are ever more apparent. Traditional teaching may have sufficed when college campuses were more ivory tower than lifeboat, educating future generations of scholars and other elites rather than trying to lift up a diverse group of students and prepare them for an increasingly complex world.

Illustration showing a shelf of lab specimen jars containing books and white-board markers. As colleges enroll students from a wider range of backgrounds, they are seeing firsthand the unintended consequences of methods such as high-stakes testing, rigid course structures, and lecture-based classes. Such traditional approaches to teaching, reformers argue, disproportionately set up students from disadvantaged backgrounds to flounder or fail. Active learning and other evidence-based practices, such as building more small assignments, or scaffolding, into the syllabus, have been shown to close those performance gaps and help all students succeed.

The problems go beyond ones of equity. Research has shown that in fields like STEM, traditional teaching can be ineffective at helping students understand complex concepts and develop problem-solving skills. Struggling students often decide early on that science and engineering are not for them.

In fact, one of the inspirations for the Harvard study was earlier work done by Carl Wieman, a Nobel Prize-winning physicist and evangelist for active learning, who has long advocated for programs that help transform science education. His former student Louis Deslauriers, now director of science teaching and learning in the faculty of arts and sciences at Harvard, and one of the authors of the study, had written a high-profile active-learning study with Wieman 10 years ago. Yet, as he and the other physics instructors noted in their introduction to the 2019

study, most STEM instructors continue to use traditional teaching methods in large introductory courses.

Why? One reason, Deslauriers says, is that they have trouble imagining why new techniques would be necessary. Whenever he would try to talk to his colleagues about what the research on teaching showed, “it would always come down to, Hey, when I was a student, traditional lecturing worked for me.”

Part of the uncertainty about research on teaching and learning stems from how it is defined. What is it, exactly? Lab experiments on how the brain works? Studies of student behavior? Experiments with teaching styles and course structure? Or perhaps a more philosophical analysis of what it means to become an expert in a discipline or a new way of thinking? The answer, in short, is all of the above.

To education researchers the terms “science of learning” and “scholarship of teaching” mean two different things. The latter term was popularized by Ernest L. Boyer in his influential 1990 book, *Scholarship Reconsidered: Priorities of the Professoriate*. Boyer, who was president of the Carnegie Foundation for the Advancement of Teaching, argued that teaching, carefully considered, is a form of scholarship and should be recognized as such.

Boyer’s call to elevate the value of teaching helped open the floodgates for faculty members to begin examining their work in the classroom, says Regan A.R. Gurung, associate vice provost and executive director of the Center for Teaching and Learning at Oregon State University. Early scholarship was typically descriptive, focusing on what professors had learned over time about their experiences as teachers.

Since then, scholarship on teaching and learning, or SoTL as it is commonly called, has become more sophisticated, complete with controls, statistical analyses, and quantitative measures of learning, says Gurung, who has written extensively about the evolution of the field. A subset, mostly found in STEM, is known as discipline-based education research, and focuses on the challenges of teaching, say, certain chemistry or physics concepts. Many disciplines now publish journals related to teaching, in which such studies appear. And more colleges are giving grants and other support to faculty members who want to do research on their own teaching.

In recent years, a new strand of research has focused on analytics — mining the data that can be found in learning-management systems and institutional research offices to ask very specific questions, such as: How does the amount of time a student spends watching video lessons or doing online reading correlate to grades? One of the potential benefits of this form of research is that it can be scaled up, looking at large numbers of courses in an institution, or longitudinally, to see how students’ performance in a prerequisite affected their ability to succeed in the more advanced course.

The “science of learning,” by contrast, most often describes the work of researchers in fields like cognitive psychology and neurology, who run lab- or classroom-based experiments on how the brain works, and how that relates to learning.

Some of the earliest, and most familiar, research of this kind involves motivation and memory. Many studies have shown, for example, that people remember things longer if they space out their learning sessions and test themselves at regular intervals rather than cramming the night before a test. Another common finding is that people make stronger connections among concepts if they review earlier ideas as they learn new ones instead of learning in discrete segments.

In 2014 the book *Make It Stick: The Science of Successful Learning* was published, eventually selling more than 600,000 copies. Mark McDaniel, one of its authors and director of the Center for Integrative Research on Cognition, Learning, and Education at Washington University in St. Louis, credits the book’s appeal to the way it translated experimental research into classroom practices, something that was lacking in the scholarship at that time.

Since then, research on learning has branched out to include the study of how emotion and environment can affect a person’s ability to learn. As colleges grapple with how to raise retention and graduation rates among struggling students, researchers have homed in on questions like: How does a student’s self-efficacy or sense of belonging correlate with academic success? How can you foster curiosity in your classroom? How does trauma affect the brain and the ability to learn?

Books such as *The Spark of Learning: Energizing the College Classroom With the Science of Emotion*, by Sarah Rose Cavanagh, and *How Humans Learn: The Science and Stories Behind Effective College Teaching*, by Joshua R. Eyler, embody this trend.

Many professors are open to using evidence-based teaching practices, notes Eyler, director of faculty development at the University of Mississippi, but would benefit from understanding the science behind them. What, for example, makes peer learning an effective technique? What do cognitive science, evolutionary biology, and neuroscience tell us about how traits such as curiosity and authenticity increase a person’s ability to learn?

Given all these strands of research and scholarship on teaching and learning, it’s not surprising that your average professor might feel intrigued yet overwhelmed. Much like trying to evaluate studies of diet, nutrition, and exercise, faculty members can struggle to determine what research is relevant for them.

Some of what works is dependent on a scholar’s discipline and teaching demands. What’s needed to engage a student in an introductory science course is different from what makes a history seminar run well. But there are also profound differences of opinion over some fundamental questions. Among them: What constitutes good evidence? How do you define learning?

Cavanagh, the author of *The Spark of Learning* and an associate professor of practice in psychology at Simmons University, in Boston, recalls an incident from a workshop about her research. She usually finds a receptive audience, often with other STEM professors who may be interested in her scholarship. In this instance she was talking to a group of humanities professors participating in a yearlong examination of the social and emotional aspects of learning. She had begun talking about how, if learning is the retention and retrieval of information and the development of new skills, then emotion may be the best route through which to engage students.

One of the professors interrupted her: Learning, he said bluntly, is not the same as remembering. Realizing the humanities professors might be operating within a completely different frame, Cavanagh moved the conversation toward a broader discussion of the role of emotion in learning.

The divide often comes down to this question: Can you measure learning? If you don't believe you can, in a quantitative way, Cavanagh says, "then you're never going to believe a research study that shows pedagogical technique XYZ boosted exam scores."

While describing the divide as a disciplinary one would oversimplify it, many humanities professors would argue that learning is a process of transformation. They are happy to study their teaching, but their scholarship is more reflective than quantitative. And they challenge their peers to take a deeper, more nuanced, look at what's happening in and around college classrooms.

This "methodological saber rattling," Gurung says, is tough. "So many of us will scoff, and rightly so, about a 30-person study that has not been replicated. And a lot of folks in the humanities will say: 'What's all this replication stuff? Let's examine my group of 30 students.' There's a lot of power in that."

Robin DeRosa, director of the Open Teaching & Learning Collaborative at Plymouth State University, in New Hampshire, suggests two other reasons that some faculty members may be skeptical of studies that rely on measurement. One is the underlying assumption that only what can be measured is relevant. Yes, collecting data is important and valuable. "But anyone who works in education with actual humans knows that data only tell small glimpses of the story," she says. "A metric cannot tell you if a student's mom died while she was taking an English course, or whether they are on the [autism] spectrum."

Illustration showing a shelf of lab specimen jars containing a clock and a lightbulb.

Professors may also be skeptical of the messaging that comes with some of this research, particularly if it's used to support a single tool or strategy. "Because higher education is in crisis now, we're very solutions oriented, we're very data driven," DeRosa notes. That can cause college leaders to think that one initiative or approach can help fix a big institutional problem,

such as a 45-percent graduation rate. “That’s a really naïve way to think about teaching. And it also does damage to the faculty.”

Disagreement exists even among scholars who focus on more quantitative research. Can a study of a single intervention in a single course, for example, say much of anything? Maybe not to anyone except instructors who teach similar courses. Are all the controls set up correctly? It’s hard to know, if you haven’t been trained in education research.

One of the reasons the physics instructors at Harvard pursued their study, in fact, was because they were troubled by the lack of quality controls in much of the work that came before them. That continues to be a challenge. “When I do research I get super excited by the titles of papers, but when I click on them and start reading the abstract it’s such a narrow, specific context and they don’t control for anything,” says Kelly Miller, an associate senior lecturer in applied physics. “It doesn’t really shed any light on the actual issues. I would say the vast majority of studies are like that.”

Some researchers are advocating for more rigor in the training of faculty members who want to do this work in their classrooms, and in the design of teaching experiments. One of the more recent innovations is a project called Many Classes, which involves a network of faculty members studying the same teaching challenges. It is a model that could represent the future of certain types of education research, says Ben Motz, who runs the project and directs the eLearning Research and Practice Lab at Indiana University at Bloomington.

The Many Classes project recruits instructors across a variety of institutions and in different disciplines to test out an intervention, giving researchers a large and diverse sample. Its first study asked a common question: Does it matter when you give students feedback on their work? It found no difference in student performance between those who had received immediate feedback from instructors and those for whom it was delayed.

Faculty developers, whose job it is to translate education research for their colleagues in the classroom, say that it often takes years and myriad experiments to draw broad lessons. That can make the research tricky to communicate.

“It’s hard for faculty to understand sometimes that the science of teaching and learning is built on lots and lots of smaller studies that give us this broader picture,” says Lindsay Wheeler, assistant director of STEM-education initiatives at the University of Virginia’s teaching center, who has studied what prevents faculty members from changing the way they teach.

Active learning broke through the noise thanks to a 2014 meta-analysis of 225 studies of STEM courses, which found that active learning increased grades and reduced failure rates, compared with lecture-style teaching.

It's easy to dismiss any one study, in other words. But collectively many point to a cohesive set of practices that improve learning.

Another problem hampers the classroom adoption of research on teaching. What feels right to students — and some professors — is not necessarily what serves them best. Active learning, as demonstrated by the Harvard study, is one such example. In their analysis, the researchers suggest that faculty members explain in advance to students why strategies such as group work will help them understand the material better, even if it sometimes feels far more difficult and less satisfying. That may increase students' willingness to try new things.

Anne Cleary, a psychology professor at Colorado State University who studies human memory, says there's a term for these kinds of learning strategies: desirable difficulties. They require a lot of effort on the part of the student, but they're necessary for learning that sticks. Yet how do you get students to break bad habits?

"I can still remember having this list of vocabulary words as a kid and sitting at my parents' dining-room table and repeating them over and over," she says. "Now I know it's one of the least effective strategies for learning. But when I ask students every semester how many think it's useful and how many do that, a large number raise their hands."

Cleary is among those professors trying to tackle that challenge with strategic interventions. Through an elective called the Science of Learning she hopes that if students read the research on memory and learning they will adopt better strategies. These desirable difficulties include strategies like testing yourself regularly on what you've learned, rather than reading the same passage over and over with a highlighter in hand. "What we're teaching people doesn't feel good," Cleary admits. And the techniques require continual practice to be effective. "It's a horrible sales pitch."

Cleary also helps other faculty members figure out how to incorporate these strategies in their teaching. Students tend not to like, say, weekly quizzes. And professors often don't want to stop in the middle of a lecture to ask students to jot down what they've learned so far. It makes Cleary uncomfortable, too. "It feels like I'm not doing anything. I'm just standing there," she says. "I should be cramming more content into my lecture."

Place all of these disagreements, uncertainties, and challenges within the structures and systems of higher education, and it becomes even clearer why research on teaching and learning has made limited inroads into the classroom. Tenured and tenure-track faculty members are under tremendous pressure to manage multiple responsibilities, including research in their own fields and service work, leaving little time to catch up on the latest study on, say, peer learning.

Contingent instructors, many of whom are in charge of large introductory courses that are extremely challenging to teach, are not compensated for the additional time it would take to sort

through much of the research on these courses. Even committing to something more than a single workshop can seem like too heavy a burden.

Gurung, a professor of psychology, has been tracking academics' attitudes toward research on teaching and learning through the years. Surveys from 2008 and 2017, he says, demonstrate a growing interest across disciplines in conducting this kind of scholarship, with faculty members in psychology leading the way. But many professors still report a lack of institutional support for the work.

Higher education also creates few incentives for faculty members to explore scholarship on teaching and learning. Tenure and promotion policies rarely reward, or even recognize, the hidden work it takes to improve one's teaching. Departments routinely rely on student course evaluations without looking at how much time a faculty member might spend trying out new teaching strategies, taking workshops through the campus teaching center, or reading the latest education research in their discipline.

Given the de-emphasis on professional development, says E. Shelley Reid, director of the Stearns Center for Teaching and Learning, at George Mason University, it's no wonder that few professors want to take risks with their teaching. "It's not like doing research in the lab and there are three or four people and you're expecting things to fail," Reid says. "It's a public performance every night: 'We've got this Broadway show. Should we tinker with it mid-run? No.'"

Mix those structural challenges into the broader culture of academe, where a stellar record of research is often held in higher regard than a reputation for excellent teaching, and it's easy to see why so many professors are unaware of the scholarship on teaching.

As early as graduate school, the message is clear. Most Ph.D. programs devote nearly all of their time training students to do research, the implication being that disciplinary expertise is all that's needed to be effective in the classroom.

"Being a good teacher isn't rewarded in the academy," says Lindsay Masland, an associate professor of psychology at Appalachian State University, in North Carolina. "Why would they know about this research? Why should they?"

Academics who might want to study their own teaching could also feel discouraged from doing so. Masland recalls how people in graduate school reacted when she said she was interested in the scholarship of teaching. "I got the feedback, 'You're too smart for that.'" So she pursued a minor in statistics, she says, "to make myself seem more serious. I wouldn't have admitted that at the time, but I did. And it helped open doors."

Masland, who spends about half her time doing faculty-development work through the campus teaching center, continues to bump up against these biases. She considers them the legacy of an

era when teaching was considered women's work, while universities were the purview of men. "The academy is a place where you're expected to perform intellectualism," she notes. "And your value depends on how badass you can be intellectually. Teaching excellence doesn't feel very rock star, for whatever reason."

In 2012 the National Research Council published an influential report urging more scholars to get involved in research on teaching within their disciplines, and described how such research can help meet fundamental challenges in science and engineering education, such as improving students' conceptual understanding and problem-solving abilities.

While discipline-based education research, or DBER, has steadily grown, integrating it into departmental work has remained a challenge, researchers say. Oftentimes there's no one in a department trained to understand this research, as it draws on other fields, such as psychology and anthropology.

Short of creating new hiring lines for faculty members trained in DBER, some institutions say the solution is to offer support for professors to study and use such research. At Miami University of Ohio, Ellen J. Yeziarski, director of the Center for Teaching Excellence created a program called DBER Associates to do just that. Professors from the same discipline dive into education research with the aim of bringing more evidence-based teaching practices into the classroom.

"That transition to practice has to happen," she says. "We can blame the practitioners or we can suck it up and make it more translatable to them."

Yeziarski has brought two cohorts into the program at Miami, each tackling a teaching challenge of common concern. The physics department, for example, is rethinking an introductory course, which may require stripping out some content in order to zoom in on core concepts. "They're very much having to put a puzzle together that maybe hasn't been solved for their course," she says. But they are digging into the research on how others have measured learning of physics concepts, and which concepts are most important to learn.

Washington University is supporting randomized teaching experiments through its Center for Integrative Research on Cognition, Learning, and Education, which embeds education specialists into departments. "It's not speedy," says McDaniel, who directs the program, noting that one department spent several years studying the impact of active learning. "It's a slow process." But, he says, it's a model that other universities could adopt. "Instructors sometimes feel like they're out there on their own," he says. This program changes that dynamic.

The University of Michigan at Ann Arbor's Foundational Course Initiative tackles the problem of implementation on a broader scale. Experts from the campus Center for Research on Learning and Teaching work with departments to restructure courses to be more engaging, reduce achievement gaps among different groups of students, and develop students' critical-thinking

skills. The work on any one course stretches over several years and involves dozens of people and reams of analysis.

“If institutions are interested in promoting change, it can’t all be left to instructors’ doing their best,” says Matt Kaplan, executive director of the center. “Especially if it involves so many pieces, as a large course does.”

What might persuade more faculty members to dive into the research on teaching and learning? Teaching experts say that professors often act when they feel a gulf between what they’re doing and what they want to achieve in the classroom. The pandemic and related social-justice movements of the last couple of years have led many to re-examine their teaching, because the effect of students’ emotional states and living conditions on their ability to learn became so clear.

Studies have also shown that faculty members are more likely to try evidence-based teaching practices if they feel they have supportive colleagues and departments. Faculty learning communities can be particularly helpful, teaching experts say, because instructors meet regularly over a series of months to tackle complex challenges, often by exploring the research and experimenting with small changes to their teaching.

Reforming teaching evaluations so that they reflect the hard work of reading and reflecting on teaching scholarship is also a critical lever for change. At Appalachian State, Masland has worked with faculty members to create a rubric listing specific teaching behaviors, such as inclusive teaching, that have been backed up by research, as a motivator to try new things. “We footnoted every behavior with a series of citations. There’s a hyperlink to every study,” she says. “That changed people’s attitudes.”

Deslauriers, of Harvard, thinks the evidence will ultimately win out. “At the end of the day, faculty really care about teaching and learning,” he says. And when they become aware that their preconceived notions may be wrong, “all of a sudden these obstacles — and I’m exaggerating a bit — kind of fall by the wayside.”