Levers FOR Change

An assessment of progress on changing STEM instruction

American Association for the Advancement of Science
with support from the National Science Foundation
About AAAS

Founded in 1848, the American Association for the Advancement of Science is an international, nonprofit organization dedicated to advancing science, engineering and innovation for the benefit of all people. With more than 120,000 individual members in more than 91 countries, AAAS is the world’s largest multidisciplinary scientific society and a leading publisher of cutting-edge research through the Science family of journals. As one of the top voices for science worldwide, we spearhead initiatives in policy, international cooperation and diplomacy, STEM education, public engagement, and more. We strive to promote and defend the integrity of science and its use, provide a voice for science on societal issues, and strengthen and diversify the science and technology workforce.

More information is available at www.aaas.org.
LEVERS FOR CHANGE: An assessment of progress on changing STEM instruction

Sandra Laursen, PhD
University of Colorado Boulder

Sandra Laursen is senior research associate and co-director of Ethnography & Evaluation Research at the University of Colorado Boulder, where she leads research and evaluation studies focusing on education and career paths in STEM fields. Her recent work has addressed equity and representation of women and people of color in the sciences, professional socialization and career development of scientists, teacher professional development, and organizational change in higher education. She is also interested in inquiry- and research-based teaching and learning, and in measuring change in teaching practice. She is lead author of the book, “Undergraduate research in the sciences: Engaging students in real science,” and of a widely cited study of inquiry-based learning in college mathematics.

With Ann Austin, she developed the StratEGIC Toolkit: Strategies for Effecting Gender Equity and Institutional Change, a research-based resource for people leading institutional change to enhance the success and retention of women STEM scholars in the academy. She is a graduate of Grinnell College and has a PhD in physical chemistry from UC Berkeley.
## TABLE OF CONTENTS

Preface: Shirley Malcom, AAAS .......................... 2

1 Levers for Change in STEM Instruction ................. 6
   INTRODUCTION ................................................. 6
   Study Design ................................................. 8
   Disciplinary Reviews ....................................... 11
   Findings from the Working Meeting ....................... 12

DISCIPLINES

2 Status of Research-Based Reform in Undergraduate Life Sciences Education and Levers for Promoting Reform ................. 17
   INTRODUCTION ................................................. 17
   RESEARCH-BASED REFORMS ................................ 18
   Reform in Classroom Practices, Laboratory Experiences, and Classroom Assessment .... 18
   Table 2.1 .................................................. 19
   Reform in Instructor Interest, Beliefs, and Attitudes ............................................. 22
   Diversity and Inclusion ...................................... 27
   FACTORS AND PROCESSES .................................. 28
   Levers for Change in Life Sciences as a Whole ....................................................... 28
   Levers for Change in Academic Departments and Institutions ..................................... 31
   Levers for Change with Individual Instructors ....................................................... 33
   CONCLUSION .................................................. 36

3 Trajectories of Reform Implementation in Chemistry and Biochemistry Undergraduate Instruction ................. 40
   INTRODUCTION ................................................. 40
   RESEARCH-BASED REFORMS ................................ 42
   Classroom Practices ......................................... 42
   Professional Societies ....................................... 46
   Classroom Practices in the Laboratory ..................... 48
   Assessment ................................................... 50
   Instructor Interest, Beliefs, and Attitudes .................. 51
   Table 3.1 .................................................. 52
   Instructor Development and Preparation for Teaching .............................................. 53
   Curriculum Resources, Development, Planning and Usage ........................................ 55
   Table 3.2 .................................................. 55
   Table 3.3 .................................................. 58
   FACTORS AND PROCESSES .................................. 59
   Factors Enabling Propagation of Chemistry Reforms .............................................. 59
   Fig. 3.1: .................................................. 60
   Factors Inhibiting or Slowing Propagation .................. 61
   FUTURE DIRECTIONS ............................................. 62

4 Adoption of Research-based Instructional Strategies in Engineering and Computer Science ................. 69
   INTRODUCTION ................................................. 69
   RESEARCH-BASED REFORMS ................................ 69
   Figure 4.1 .................................................. 70
   Our Data Regarding RBIS Uptake ............................................. 71
   The Accreditation Board for Engineering and Technology ........................................ 72
   Why Do Instructors Adopt/Not Adopt RBIS? ....................................................... 72
   What Can We Do to Promote Greater Adoption of RBIS? ......................................... 73
   Where is the Future? ........................................... 74
   Diversity and Inclusion ...................................... 75
   LOOKING TO THE FUTURE ..................................... 76
   CONCLUSION .................................................. 77

5 Research-based Instructional Reform in the Geosciences: Building a Community of Practice ................. 81
   INTRODUCTION ................................................. 81
   RESEARCH-BASED REFORMS ................................ 84
   National Surveys and Observations ............................................. 84
   Table 5.1 .................................................. 86
   What do the Data Tell Us? ........................................... 86
   Table 5.2 .................................................. 87
   Figure 5.1 .................................................. 88
   What Worked and Why? ........................................... 91
   ONGOING REFORM EFFORTS .................................. 93
   Creating a Formal Geoscience Education Research Community .................................. 93
   Earth Educators’ Rendezvous (Annual Meeting) ............................................. 93
   Building Effective Resources for Geoscience Teaching ........................................... 94
I almost didn’t make it out of university intact as a STEM major. Freshman year was not easy. And while I didn’t expect it to be, I now understand that it didn’t have to be as hard as it was. Yes, I brought content and skills deficiencies from my under-resourced, segregated school in Birmingham, Alabama; but I brought strengths as well. I was highly motivated and willing to work hard. I started as pre-med and took the classic course lineup that included mathematics and introduction to chemistry, lecture and lab. At times, taking these courses felt like “hazing,” something I just needed to “get through” so that I could be accepted as a science person and achieve my career goals. As I labored through the four years of required sequences of coursework in mathematics, chemistry, physics and biology (typical for a pre-med), I had more or less difficulty in each class, depending on the amount of material I encountered, how much interest I could muster in the content, how hard it was to find the storyline, what the learning goals were, and so on.

As a person who has since spent a large part of my career supporting efforts to diversify the STEM community, it has been important to look to the research to understand the extent to which presentation of disciplinary content within courses and across the curriculum affects whether someone is successful in mastering that content and being retained in STEM. This includes the amount of content, its organization, and its design. We also know from other research that context is important, especially for students from underrepresented groups. Without that context, without building the connection between content across disciplines, it likely won’t be engaging.

The disciplinary contributions and conversations in Levers for Change provide a great deal of insight into the structure of knowledge across the STEM fields, and how that intersects with the expectations of learners and the structures of the institutions where teaching and learning take place. For example, because the geosciences tend not to be on the “must take” requirements list for almost any student in other STEM majors, the discipline has behaved in ways that reflect the expectation that it must attract students by making clear how it connects to their lives. Their methods of providing strategies and stories to develop know-how within a tightly knit community become a hallmark and a lesson to all STEM fields. At the same time, demonstrating a pathway to a career in the geosciences was recognized as an unresolved challenge in this working meeting.

Most practitioners realize that introductory courses can be critical for giving students a sense of what the field is like. Unfortunately, it is not always understood that this does NOT mean trying to force feed students EVERYTHING going on in a field. Over time, biology has been among the greatest offenders in saying too little about the core ideas of the discipline by trying to say too much! Terminology crowded out the story. The life sciences community came together and re-envisioned biology. The result was a report, Vision & Change in Undergraduate Biology Education, which stripped away the minute details and found the conceptual scaffolding around which different content could be provided. Faculty and students could go small (molecular) or big.
(ecosystem), but the same scaffold could still be used; and it still made sense, allowing student
movement across levels without loss of the storyline.

Through the *Levers* discussions it was clear that project-based strategies in engineering and
computer science are considered very compelling within and across disciplines, as is the
tradition of capstone projects. While there are attractive curricular options, the challenges
appear to be related to implementation, or as Laursen calls it, “uptake.” Enrollment and degree
data in these fields speak to another major challenge: the extent to which diversity, equity, and
inclusion are not being addressed.

Over the past decade, physics and astronomy have continued to make major contributions to
the base of research on instructional practices as well as the development of materials and
assessments. Many other disciplines have followed the lead of the Physics Education Research
(PER) community and have their own rigorous discipline-based education research base. As
noted in *Levers*, challenges in physics and astronomy have been those of implementation—
applying the research findings to the classroom and supporting widespread adoption of the best
instructional practices—and minimal efforts to support diversity and equity by creating inclusive
class environments.

Mathematics and statistics have a tightly knit community which should make changes in vision and
principles more easily accomplished. When challenged by national policy statements in the PCAST
report *Engage to Excel*, the community was able to mobilize quickly to begin to address some of
the complaints lodged against the field, especially the role of mathematics as a barrier to student
entry, retention, and success in STEM generally. But agreement at the leadership level does not
address the professional development and implementation needs at the practitioner level.

Mathematics’ dual role as a stand-alone discipline and a competency needed by all other areas
of STEM offers challenges, not unlike the challenges emerging in computer science. Can the
core infrastructure of the field lead to strategies that can accelerate implementation? Or does
the core infrastructure and tight within-discipline focus make it more challenging to reach out
for the context that the other disciplines might bring to their field?

Chemistry also enjoys the advantages of a cohesive community as well as the disadvantages
of contextual challenges as it interacts with other STEM fields. While laboratory skills are seen
as critical to the discipline, the cross-walk between what the labs provide and the content of
the lectures is often misaligned: students are left to figure out the connections on their own.
And yes, lectures predominate, despite research demonstrating the value of active learning.
Chemistry is not alone in that struggle.

As we seek to explore levers for change within each discipline, we see that the greater umbrella
of STEM itself is changing, with convergence rapidly punching holes in disciplinary walls. So, the
challenge for all is to provide the fundamentals (concepts, content, tools, ways of thinking) of
each field, while simultaneously demonstrating connections to other areas of STEM.

Beyond the silo of specific disciplines, we look for other levers for change within the institutions
themselves—what is valued and rewarded? What professional development opportunities and
support options are provided to faculty, graduate students, and post-doctoral scholars? Or we
look to signals from the disciplinary societies—who or what in our field do we celebrate? How
comfortable are we with challenging the status quo such as articulating expectations of what is core in our fields?

We are fortunate to have active communities of discipline-based education researchers to inform our work going forward; we are even more fortunate that many of these researchers are seeing value in collaboration (e.g., see the emerging STEM DBER Alliance) where they can learn from each other and tackle some of the challenges that plague us all: how do we support success by learners from underrepresented and underserved groups? How can our fields move from weeding out to cultivating talent?

The Levers discussions and discussion papers have given us an opportunity to reflect on each field’s strengths and challenges, knowledge and gaps. While each disciplinary group had positive stories to bring, it was the emergence and articulation of shared concerns that hold significant promise for improvements across all fields: what is needed to support and promote implementation; what can be done better in the organization of our content, curriculum, and courses to support all learners and to promote diversity and inclusion; how do we move our different fields toward each other in support of the emerging STEM focus on convergence?

While the Levers discussion may have been our first such conversation to synthesize and share research on the status of our individual disciplines it should not be the last.

I am reminded of the folk story, Stone Soup. In the tale, a weary and hungry traveler was passing through a village where the people living there were experiencing hard times. The villagers were cold and inhospitable to the stranger and even distant to each other. In the story the traveler proceeded to build a fire, place a huge pot of water over the fire, and bring out a ladle, knife and stone as though preparing to cook. The curiosity of the villagers caused them to approach and ask the traveler what he was doing, “Making stone soup,” he replied. Assuring them that it was really good, he offered to share once it was done. As he stirred the pot the crowd continued to grow. “This would be even better with a few potatoes,” he suggested. One of the villagers volunteered a few potatoes which the stranger accepted with gratitude, peeled, cut up and put into the pot. As you can predict, the traveler kept suggesting other things that would improve the soup, and villagers kept offering things (carrots, onions, salt, pepper, peas from the garden and more) that were added to the pot. It was the best soup ever, not because of the stone but because of what the villagers had shared. Our intellectual stone soup has given us a clearer way forward in our separate, collective, and collaborative efforts to improve the quality of undergraduate STEM education.

Shirley M. Malcolm
INTRODUCTION

To meet the demands of a global economy and foster technological innovation, the United States needs more well prepared and diverse workers in science, technology, engineering and mathematics (STEM) fields (PCAST, 2012). National studies reveal racial and ethnic disparities in science literacy, as well as in educational achievement, employment, and health outcomes that depend on STEM education (Allum, Besley, Gomez & Brunton-Smith, 2018). All Americans should have equitable opportunities to enter the high-paying, high-status, and high-employment jobs typical of STEM careers, and to learn, enjoy, and use science to make informed decisions in everyday life, in the voting booth, and in their communities (Rutherford & Ahlgren, 1989; Snow & Dibner, 2016). Access to STEM learning opportunities begins in childhood and requires well-prepared preK-12 and informal educators to teach and inspire young people in mathematics and science (PCAST, 2010). High-quality STEM education for all undergraduates is essential to achieving all of these national goals.

A large and ever-growing body of education research demonstrates that pedagogical approaches that foster active and collaborative learning can enhance student learning, attitudes, and persistence in STEM educational paths (for reviews, see Fairweather, 2008; Freeman et al., 2014; Froyd, 2008; Ruiz-Primo et al., 2011; Springer, Stanne & Donovan, 1999; also Braxton, Milem & Sullivan, 2000; Colbeck, Cabrera & Terenzini, 2001; Hake, 1998; Laursen, Hassi, Kogan & Weston, 2014; Madsen, McKagan & Sayre, 2015; Von Korff et al., 2016; Wang, Sun, Lee & Wagner, 2017). Yet most students do not experience these engaging pedagogies. Indeed, students from underrepresented racial and ethnic groups, as well as low-income and first-generation college students, are more likely to benefit, yet least likely to experience them (Kuh, 2008; Museus, Palmer, Davis & Maramba, 2011). Women also benefit disproportionately (Kogan & Laursen, 2014; Laursen et al., 2014; Lorenzo, Crouch & Mazur, 2006). Policymakers view improving instruction as a “best bet” (Singer, 2013, p. 768) and as the “lowest-cost, fastest policy option to providing the STEM professionals that the nation needs” (PCAST, 2012, p. i).

If active and student-centered pedagogies are an important solution to national challenges in STEM education and workforce development, how can we ensure that undergraduates routinely experience them in their STEM classes? A decade ago, Fairweather (2008) reviewed the literature on promising practices in STEM undergraduate instruction and concluded that the problem was not a lack of knowledge about which teaching practices were effective, but rather insufficient use of these practices:

_The key to improving STEM undergraduate education lies in getting the majority of STEM faculty members to use more effective pedagogical techniques than is now the norm in these disciplines._ (p. 13) ...[M]ore effort needs to be expended on strategies to
promote the adoption and implementation of STEM reforms rather than on assessing the outcomes of these reforms. ... The problem in STEM education lies less in not knowing what works and more in getting people to use proven techniques (p. 28).

Of course, efforts to get people to use effective, research-based teaching methods must attend to what is known about the internal and external barriers that hinder instructors from doing so. Most PhD-trained academics have little formal training in teaching and learning. Their mental models of teaching and learning may emphasize transmitting accurate information rather than helping learners construct and connect ideas; their beliefs about students may emphasize what students can’t do rather than what they can (Henderson & Dancy, 2007; Lund & Stains, 2015). They may not believe students will voluntarily or capably engage in classroom practices that demand more mental effort, and they may not be equipped with ways to respond if they do meet resistance (Finelli, et al., 2018; Seidel & Tanner, 2017).

Academic values and identities developed in disciplinary socialization, reinforced by academic reward systems, tend to deter faculty from focusing on teaching at the expense (real or perceived) of research (Austin, 2011; Brownell & Tanner, 2012; Marsh & Hattie, 2002; Seymour, 2002). And even beliefs consistent with learner-centered teaching approaches are often insufficient to move instructors to adopt these approaches. Situational constraints also shape instructors’ choices, such as requirements to coordinate coverage and pacing in multi-section courses; physical spaces that limit student peer-to-peer interaction; implementation costs, particularly for laboratory reforms; and departmental norms of teaching (Borrego, Froyd & Hall, 2010; Dancy & Henderson, 2010; Henderson & Dancy, 2007; Hora & Anderson, 2012; Lund & Stains, 2015; Parker, Adedokun & Weaver, 2016; Walczyk, Ramsey & Zha, 2007). Thus, it is crucial that we learn how to lower these barriers and promote adoption of effective evidence-based teaching practices (Colbeck, 2002; DeHaan, 2005; Henderson, Beach & Finkelstein, 2011; Hora & Ferrare, 2013; Seymour & De Welde, 2016).

We consider this problem in terms of a metaphor, “levers for change.” Why levers? In mechanics, a lever is a simple machine used to move an object at one location by applying a force somewhere else. By working at a distance, a lever acts to magnify the applied force. Metaphorically, then, a lever is a means to achieving an end, a method of persuading
or causing something to happen. When we try something and see that it is working, we have gained leverage on the problem. Here we consider how systems, structures, and cultures can work as levers to accomplish change in STEM instruction, recognizing the complex and multi-level nature of the environment in which such levers operate (Austin, 1998, 2011). For example, studies of instructors have shown that one commonly cited reason instructors do not use research-based instructional strategies is a lack of time to prepare new materials or implement new strategies (Dancy & Henderson, 2010; Henderson & Dancy, 2007; Lund & Stains, 2015). But what makes instructors feel they don’t have time for such work? What changes to cultures and structures might give instructors a sense of having “more time” by redirecting their priorities or providing professional sanction for their efforts to learn, apply and master research-based instruction? It is such levers that we consider in this study.

STUDY DESIGN

In this section we describe the goals and design of the study that informed this report, the research questions, and the sources of information used.

Goals

We set out to assess the state of reform in STEM undergraduate instruction and to identify effective levers for change on STEM undergraduate teaching and learning both within and across six STEM disciplinary clusters. We sought community-based answers to the following questions:

▲ What is the current state of research-based reform in undergraduate instruction within these six clusters of STEM disciplines?

▲ How did each arrive there? What levers for change—activities, events, influences, movements, groups, documents, contexts—have been important in reaching this state? And how are these levers similar or different by discipline?

▲ What provides evidence for these trajectories of change, and why?

▲ What can be learned from this evidence about how to expand and deepen the impact of these changes in the next decade?

We explored these questions through both scholarly and practical perspectives, drawing on the research and evaluation literature on STEM instructional reform in higher education and on knowledge derived from practical experience in conducting, leading and observing such reforms. The study, therefore, had two components. First, to gather and assess knowledge from the literature, we commissioned scholarly essays by researchers who study instruction and change in each of the six disciplinary clusters. Then, to gather and assess knowledge from practitioners, we convened a working meeting of about fifty people involved in a wide swath of instructional change activities at a diverse array of institutions. This combined approach was important because it seemed likely that practitioners could offer observations and conjectures that have not yet been examined in the literature. More details about each of these components are provided below.

The resulting report offers a snapshot of the current state of reform within and across STEM disciplines. We then consider what this current state of STEM instruction suggests about the opportunities for the next decade or two, using what has been learned so far to suggest strategies for the future. By highlighting these opportunities, we hope the report will be useful for stakeholders in STEM
higher education, including researchers, practitioners, and change agents working in institutional, disciplinary, and cross-institutional arenas.

**Scope and Assumptions**
This project was organized using a disciplinary approach in explicit recognition that disciplines may have taken different trajectories to change. People working in different disciplinary settings may have self-organized differently, made use of different assets, or encountered different barriers, and these differences may depend on their disciplinary practices, norms, and cultures. Our disciplinary approach also recognizes that some knowledge of these issues comes from the work of scholars in discipline-based education research (DBER) who have studied teaching or evaluated approaches to instructional change in their own fields. Thus, five of the six disciplinary clusters considered here parallel the fields covered in the National Academies’ study of discipline-based research in education (Singer, Nelson & Schweingruber, 2012). This project added the mathematical sciences, including statistics, and combined astronomy with physics because of their close alignment in undergraduate education and DBER scholarship. The project engaged people with scholarly and practical expertise in six clusters of disciplines:

- **LIFE SCIENCES**
- **CHEMISTRY AND BIOCHEMISTRY**
- **ENGINEERING AND COMPUTER SCIENCE**
- **GEOSCIENCES**
- **MATHEMATICAL SCIENCES, INCLUDING STATISTICS**
- **PHYSICS AND ASTRONOMY**

In addition, participants from higher education and funding agencies contributed their knowledge of cross-disciplinary change efforts and helped to generate comparative perspectives.

To create boundaries for the project and to communicate those parameters clearly to participants in the working meeting, we identified five working premises and asked participants to accept these premises in order to maintain focus on the questions at hand. While other assumptions and boundaries would be defensible, we found these premises helpful in shaping the reviews and working discussions, even as we acknowledged their limitations.

1. **There is sufficient evidence from education research** in and across the disciplines to indicate that active-learning experiences are good for students and support their learning, attitudes, sense of belonging, and persistence in STEM. (We know that ongoing studies will further detail these benefits and how they vary among different student groups and settings.)

2. **We take a broad view of what strategies count as active learning.** (We acknowledge that there is more to learn from comparing and contrasting strategies and their affordances and limitations for students and instructors. We also recognize variation in the extent of evidence behind different strategies.)

3. **We focus on classroom instruction provided by college instructors in STEM disciplines.** (We recognize that co-curricular and extracurricular experiences also matter for students.)

4. **Instructors’ adoption of active teaching strategies is a critical step in ensuring that all students experience the benefits of active learning in their STEM courses.** (We include not only initial adoption but skillful and sustained use of these strategies, which takes time to develop.)
We use the term “evidence” inclusively, considering scholarly evidence from research and evaluation, but also experience and observations from the field. (We recognize that the state of knowledge in this domain is evolving.)

We also define some common terms used throughout the report. First, the term research-based instructional strategies (RBIS) is used to designate the set of active teaching and learning practices that support improved student learning (premises 1 and 2). In general, such active, collaborative, and student-engaging strategies support learning, independent of discipline (Kuh, 2008; Pascarella & Terenzini, 2005; see also Fairweather, 2008). However, the form of specific RBIS may vary among disciplines, as may the extent to which specific methods are supported by published scholarly research.

Second, the term instructor is used to describe all people with undergraduate teaching duties, focusing on their educational role, not their job title (premises 3 and 4). Instructors may include tenure-stream (pre-tenure and tenured) and non-tenure-track faculty with or without job security, as well as graduate teaching assistants and even undergraduate learning assistants, all of whom help to enact STEM instruction for undergraduate students.

Finally, we use the shorthand uptake to reference the overall degree of instructors’ adoption and use of RBIS. Uptake is used here to reference the net state of reform toward the use of RBIS, not to describe the process or situation of any single instructor. Drawing on its use in chemistry and cooking, the word is intended to evoke both change and retention, as in a plant’s uptake of nutrients from the soil, a person’s uptake of vitamins from food, or uptake of flavors by butter added to a savory dish.

Data Sources
The study includes elements to probe both scholarly and experiential knowledge about levers for change in the six disciplinary clusters (premise 5). Essays commissioned from disciplinary reviewers summarize the scholarly knowledge available in each discipline, and a working meeting was designed to extract and synthesize experiential knowledge. These two components are described in detail below so that readers understand the source and basis of the claims made in the scholarly reviews (Chapters 2-7) and the practice-based findings (Chapter 8) of this report.
While originally conceived primarily as working documents, the six disciplinary reviews of the research proved to be thoughtfully prepared, judicious syntheses of the literature. We include them in full as Chapters 2-7 of this report, and we describe here the scope of the review task that reviewers undertook. In particular, these scholars were not asked to conduct a comprehensive literature review, which was beyond the scope of the time and resources provided. Rather, they prepared a scholarly essay drawing upon the literature to offer answers to two main questions about the extent of research-based reform in their field, and the factors and processes that have influenced the changes they identify. The following questions helped to structure their reviews and ensure that each review provided similar coverage of the issues.

1. **To what extent, and in what ways, is research-based reform in undergraduate STEM instruction occurring in your own STEM discipline? How do you know?**
   - What evidence from research and practice is available about the nature and extent of implementation of research-based reforms in STEM instruction? What evidence is missing or inadequate?
   - What does that evidence tell us?
   - What changes have been observed or reported, and in what domains? Which types of change seem to be more established, and which less so?

2. **What factors and processes have influenced the depth, breadth, and impact of these changes?**
   - What factors—events, documents, people, movements, contextual circumstances—have influenced, in positive or negative ways, the depth, breadth, and impact of these changes? For example, in the life sciences, the Vision and Change report (AAAS, 2011) is thought to be one such influence. The influences identified may be discipline-specific or broader.
   - What lessons can be learned from the available evidence about ways to measure or monitor the nature and extent of research-based reform in undergraduate STEM instruction? For example, Foote et al. (2014) offer ideas for how to measure the spread of a particular research-based reform practice within a discipline.
   - What do we not know that is important about the extent of implementation of research-based reform in STEM instruction or about factors affecting such changes?

We asked reviewers to use the scholarly literature to support their claims where possible but recognized the limitations of available literature. Knowing that generalizable studies would be in limited supply, the evidence includes examples and case histories. Finally, we invited authors to draw on less-formal sources if they wished: information from professional societies, conference programs, disciplinary groups, evaluations or records of funded projects, websites, and wise colleagues. The final essays draw on scholarly evidence, the authors’ collegial networks, and their considered judgment.

To define the scope of the reviews, we asked reviewers to consider indicators of reform at different levels—individual instructors; organizational units such as departments, colleges and institutions; and the discipline as a whole—and to address several domains.
where change might occur, at different paces. Authors could include other domains important in their discipline, but were asked to discuss four core domains explicitly:

- Classroom practices and assessment;
- Instructor interest, beliefs, and attitudes related to reform-based teaching;
- Instructor development and preparation for teaching; and
- Attention to diversity and inclusion in undergraduate STEM instruction.

For each review, we provided guidelines as to what sub-disciplines to include. For example, we asked both reviewers for biology and chemistry to include information on biochemistry if they found it in their reading; it was formally assigned to the chemistry working group at the meeting.

As for participants in the working meeting itself, we asked reviewers to assume that the outcomes of research-based reforms in instruction are positive for STEM undergraduates’ learning, attitudes, persistence, and/or preparation for work or further education. That is, they were not asked to make the case that these reforms, if implemented, will have positive impacts. Rather, the reviews take as given that there is sufficient evidence about student outcomes from sources such as meta-analyses (e.g., Freeman et al., 2014; Ruiz-Primo et al., 2011) to make a “golden spike” argument to connect sound implementation of such reforms to improved student outcomes (Brown Urban & Trochim, 2009). Readers should recognize this as a design choice for this project, not evidence of any lack of nuance in reviewers’ understanding.

**FINDINGS FROM THE WORKING MEETING**

The working meeting convened 52 participants on May 7-8, 2018, at the Howard Hughes Medical Institute (HHMI) campus in Bethesda, MD. Several people from HHMI and the National Science Foundation also participated. Working groups of 7-9 people convened in each disciplinary cluster, with participants representing diverse perspectives across their discipline: STEM education researchers interested in instruction and change, faculty developers, department heads and deans, emerging and experienced instructional leaders, and people involved in change initiatives organized by disciplinary societies, higher education organizations, and funding agencies. The meeting design was informed by a prior meeting, also convened by the American Association for the Advancement of Science, that examined the state of knowledge on methods for characterizing and describing teaching (AAAS, 2013).

While most participants had worked within a particular STEM discipline, one risk of a disciplinary approach to the meeting was the potential to miss important ideas that cut across STEM or that were not discipline specific. To mitigate that risk, we included some people who had studied STEM instructional change more generally or who worked on initiatives that involved multiple disciplines; they joined various working groups over the course of the meeting. We also used some sessions to mix people by discipline and to extract comparative observations.
As a group, the meeting participants included people from two-year, bachelors, masters, and doctoral degree-granting institutions and minority-serving institutions; people with long experience leading or observing reform and emerging leaders who will continue such work; and people with diverse career and life experiences. Most participants were from U.S. institutions; one Canadian and one Australian institution were also represented. We asked participants to consider the guiding questions from the points of view of their multiple roles, and to represent their projects and their close collaborators as well as their own perspectives. A list of meeting participants is provided as Appendix A.

Two types of leaders helped guide each disciplinary working group. The reviewers were discipline-based education scholars with expertise on instructional change in their field. Each reviewer or review team prepared a white paper, based on the literature, to stimulate meeting participants’ thinking in advance of the meeting and to provide a common basis of knowledge about available evidence from the literature, as discussed above. The reviews also helped shape the discussion topics and guiding questions for the working meeting. The conveners served as facilitators for working group breakout sessions and helped gather participant input on the discussion questions. To prepare for the meeting, participants were asked to read the disciplinary review for their own field, plus one additional review. These assignments were distributed so that at least one person in each discipline-based working group had read each review from the other fields, ensuring that the group as a whole had access to all the ideas thus introduced. The six review teams also gave short presentations of highlights from their reviews to kick off the meeting. A list of reviewers and conveners is provided as Appendix B.

Through presentations, structured whole-group discussions, and breakout sessions in different configurations, the working meeting was designed to explore and to capture participants’ assessment of the nature, depth and breadth of reforms in undergraduate STEM instruction, their sense of the key factors in advancing those reforms, and their views of the opportunities and challenges looking forward. Guiding questions provided structure for each segment of the meeting. Groups recorded their discussions of these questions in shared documents and reported important ideas periodically to the full group through posters and verbal reports.

This material, together with the disciplinary reviews, the meeting organizers’ notes, and transcripts based on audio-recordings of group discussions, was used for a careful qualitative analysis that is the basis of the current synthesis report from the working meeting. While we did not ask disciplinary groups to come to consensus, there was in fact a good deal of commonality in their conclusions. In general, we found the groups’ notes to be self-critical rather than self-lauding; they drew on their own experience but also brought familiarity with a wide range of change efforts in their field. While 52 people cannot represent the entire nation, we have some confidence that the findings based on their discussions are not idiosyncratic. The agenda for the meeting is provided as Appendix C, and findings based on the expert groups’ discussions are reported in Chapter 8.
REFERENCES CITED


President’s Council of Advisors on Science and Technology (PCAST) (2010). Report to the President. Prepare and inspire: K-12 education in science, technology, engineering, and math (STEM) for America’s future. Washington, DC: Executive Office of the President, President’s Council of Advisors on Science and Technology.
President’s Council of Advisors on Science and Technology (PCAST) (2012). Report to the President. Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. Washington, DC: Executive Office of the President, President’s Council of Advisors on Science and Technology.
INTRODUCTION

This chapter addresses two questions about the uptake of research-based instruction strategies (RBIS) in undergraduate life sciences education as set forth in the overall study design:

1. To what extent, and in what ways, is research-based reform occurring in undergraduate life sciences education?

2. What factors and processes have influenced the depth, breadth, and impact of these changes?

This paper does not describe the evidence that research-based reforms are positive for undergraduates’ outcomes, nor does it distinguish in any systematic way between different types of RBIS. It also is not exhaustive. Rather, I have aimed to highlight programs, research, and evidence that I think illustrate our current state of knowledge regarding the questions above. Note that these questions are specific to life sciences. Therefore, I am not drawing on the vast research that addresses these questions in other STEM disciplines. Instead, I focus narrowly on work investigating research-based reform in the life sciences to highlight what we know that may be unique to the life sciences and what we have learned in the context of life sciences that may be relevant to other STEM disciplines.

I have emphasized areas where existing evidence is lacking or insufficient to help us confidently answer these questions, which was is often the case. I have drawn on peer-reviewed research, but other sources of information have also been crucial, including the individuals acknowledged above. Lastly, I drew on my own informed perspective about the status of what we know about levers for change in undergraduate life sciences education. I expect that this adds value to the paper, while also revealing my own short-sightedness.
This first section describes research-based reforms as implemented in different learning environments, including classrooms and laboratories, as well as what is known about new approaches to assessment. It also reviews different professional development programs for instructors at all levels, which can impact how these reforms are introduced, and whether or not departments and instructors in the life sciences are actively committed to diversity and inclusion in the classroom.

**REFORM IN CLASSROOM PRACTICES, LABORATORY EXPERIENCES, AND CLASSROOM ASSESSMENT**

Life sciences instructors have been called upon to integrate more research-based and student-centered instructional practices, to create new high-quality opportunities for students to apply the process of science and build scientific thinking skills, and to help students build skills for collaborative interdisciplinary work (AAAS, 2011). This report, *Vision & Change in Undergraduate Biology Education*, has been a catalyst and guide particular to the life sciences over the past years (Austin, 2018). The paragraphs below summarize what we currently know about the use of RBIS in classrooms, undergraduate research experiences, and classroom assessment.

**Classroom Practices**

Overall, life sciences instructors report extensive use of lecturing and moderate levels of RBIS use. Use varies considerably across different RBIS and remains low for some strategies. Additionally, only a small proportion of life sciences instructors seem to be extensively relying on RBIS in their teaching.

The Higher Education Research Institute’s (HERI) Faculty Survey has a large scope, surveying 1000 or more life sciences instructors every three to four years (Eagan, 2016). Data collected in 2014 are summarized in Table 2.1 and can be compared to prior years to learn about change over time. Over 80% of life sciences instructors reported extensive use of lecturing in all or most undergraduate classes they taught in 2004; this dropped to slightly less than 70% in 2014. The use of class discussions has remained relatively stable since 2007. In contrast, more instructors reported using electronic quizzes with immediate feedback (up from 9%), student inquiry to drive learning (up from 38%), and group projects (up from 29%) in all or most classes in 2014 compared to 2007 (Table 2.1; Eagan, 2016).

When asked about more specific RBIS, faculty report lower levels of use. Among a random sample of 33 life sciences instructors from large institutions around the United States, 57.1% reported that they never used activities in which students use data to answer questions while working in small groups. Similarly, 49% never use clicker questions that test conceptual understanding and 43% never ask students to discuss in pairs or small groups to answer a question (Andrews, Leonard, Colgrove, & Kalinowski, 2011).

Though useful in providing insight about the extent to which RBIS have taken hold, instructor self-reported use of RBIS does not give a full picture of what is occurring in undergraduate life sciences classroom. For example, some comparisons of the practices that instructor say they are using in the classroom are not well aligned with what observers see occurring in these classrooms (e.g., Ebert-May et al., 2011). Students provide another perspective about the use of RBIS in life sciences courses. A tool called the Measurement Instrument
for Scientific Teaching (MIST) can be used to gather data about instructional practices in STEM courses from students enrolled in these courses (Durham, Knight, & Couch, 2017). Data from 87 classes (91% of which were biology classes), collected by administering the MIST to 7,767 undergraduates, showed that instructional practices that engage students in scientific practices (e.g., formulating hypotheses, critiquing experimental strategies, designing experiments, interpreting data, creating graphs, critiquing scientific literature) were less commonly used than strategies such as polling (i.e., clicker) questions, in-class activities, and whole-class discussion in which students respond to their classmates (Durham et al., 2017).

The developers of the MIST (Durham et al., 2017) have also developed another version, the MISTO, that allows comparisons among the perspectives of the students, instructor, and observers using a common set of items. Data collected from these three perspectives for 68 life sciences courses showed that the perspectives were consistent for some teaching practices, such as the use of polling and group work, but less consistent for more complex practices like engaging students in scientific practices. They are not the first researchers to find relatively strong alignment between instructor-reported practices and those reported by outside observers such as students (e.g., Andrews et al., 2011; Smith, Vinson, Smith, Lewin, & Stetzer, 2014; Owens et al., 2018). Durham and coauthors’ forthcoming work clarifies when instructor self-reported and student-reported practices are trustworthy, and when outside observers may be important for painting an accurate picture of the use of RBIS. These data also highlight the need to look more closely at the nature of the student-centered strategies in biology classrooms (i.e., what RBIS are being used, and which might be most important to student success?).

Direct observations of instructional practices are particularly valuable but are not often conducted on a large scale. Researchers made impressive headway in a 2018 study that characterized instructional practices using outside observers in over 2000 classes taught by more than 500 STEM faculty members, including 591 biology classes (Stains et al., 2018). In about half of the biology classes, the instructor lectured for more than 80% of the class. Thirty-two percent of the classes were “interactive lectures” that included some group work, and less than 20% of the classes consistently used student-centered practices (Stains et al., 2018). These data tell a slightly less optimistic story about the current status of

<table>
<thead>
<tr>
<th></th>
<th>Percent reporting use in all or most classes</th>
<th>Percent reporting use in no classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extensive lecturing</td>
<td>68(^a)</td>
<td>8(^a)</td>
</tr>
<tr>
<td>Class discussions</td>
<td>70(^a)</td>
<td>5(^a)</td>
</tr>
<tr>
<td>Electronic quizzes</td>
<td>23.0</td>
<td>62(^a)</td>
</tr>
<tr>
<td>Student inquiry</td>
<td>55.7</td>
<td>10(^a)</td>
</tr>
<tr>
<td>Techniques to create</td>
<td>43(^a)</td>
<td>28.3</td>
</tr>
<tr>
<td>Group projects</td>
<td>50.4</td>
<td>18(^a)</td>
</tr>
</tbody>
</table>

\(^a\) These numbers are visually estimated from graphs and are, therefore, less precise.

Table 2.1. Percent of life sciences faculty in 2014 reporting the frequency with which they use six practices in the undergraduate courses they teach (data from Eagan, 2016).
RBIS than the HERI faculty survey. At best, half of biology instructors are using student-centered practices, and only about one fifth do so extensively.

**Research-based and Inquiry-based Laboratory Instruction**

Life sciences educators agree that research experiences are a critical part of an undergraduate education and may be more impactful the earlier they occur in an undergraduate career (Elgin et al., 2016). One promising approach for engaging more students in research is the course-based undergraduate research experience (CURE). CUREs are “learning experiences in which whole classes of students address a research question or problem with unknown outcomes or solutions that are of interest to external stakeholders” (Dolan, 2016). CUREs have the potential to engage many more students, and a more diverse population of students, than the traditional, research apprenticeship model where instructors work with one or a few students within their research lab (Elgin et al., 2016). CUREs engage students in scientific practices, collaborations with peers, and communicating about science (Auchincloss et al., 2014), and may help increase inclusion and broaden participation in STEM (Bangera & Brownell, 2014).

We lack direct quantitative data about how often CUREs are being developed and used in undergraduate life sciences education, but indirect data are promising. Publications describing and investigating CUREs have increased considerably in recent years, which indicates increased attention from biology education researchers and educators (NASEM, 2017). In CBE—Life Sciences Education, a leading journal focused on undergraduate life sciences education, about 85% of articles published about CUREs have been published since 2013. A network of people and programs, called CUREnet, has been established and funded by NSF to facilitate CURE projects in biology and research about CUREs, and to broaden the diversity of faculty and students involved in CUREs. This network provides infrastructure to investigate the uptake and use of CUREs in life sciences education.

Another research-based change in undergraduate life sciences instruction is the replacement of traditional “cookbook” exercises in laboratory classes with inquiry-based instruction (Beck, Butler, & de Silva, 2014). A survey of 65 colleges and universities in 2005 found that close to 80% of institutions reported using inquiry-based approaches in biology laboratory courses (Sundberg, Armstrong, & Wischusen, 2005), compared to estimates of about 10% in 1993 (Sundberg & Armstrong, 1993). More recently, researchers have examined the quality of inquiry instruction, and found that the research base for the effectiveness of published inquiry exercises has room for improvement (Beck et al., 2014). Finally, in 2015 a working group within the Association for Biology Laboratory Education (ABLE) surveyed all of its current and past members on their current practices in laboratory education. These data provide more insight into the use of inquiry-based exercises across different types of courses. ABLE members reported that, on average, courses for non-biology majors used fewer inquiry-based activities than did courses for majors. Additionally, inquiry-based activities were more common in upper-division courses than in introductory courses (ABLE leadership, personal communication).

**Classroom Assessment**

Assessment practices are important because they can impact student behavior and thus student learning. For example, life sciences undergraduates use different studying strategies to prepare for multiple-
choice questions than they use to prepare for constructed-response questions (Stanger-Hall, 2012). Overall, research on assessment practices in undergraduate life sciences education is sparse. However, a few studies indicate that current assessment practices may fall short of holding students accountable for higher-order thinking. An analysis of over 9,713 assessment items from high-stakes assessments (e.g., quizzes and exams) in 77 introductory biology courses at 44 institutions found that 93% of items only required students to recall and understand information (i.e., knowledge and comprehension in Bloom’s taxonomy) and not to apply, analyze, synthesize, or evaluate ideas (Momsen, Long, Wyse, & Ebert-May, 2010). A similar comparison within one institution found that the cognitive level of items on biology exams was significantly lower than that of items on physics exams (Momsen, et al., 2013). The 3-Dimensional Learning Assessment Protocol (3D-LAP) provides a new framework for analyzing assessments in undergraduate biology education that is well-aligned with Vision & Change (Laverty et al., 2016). This tool has yet to be used on a large sample of biology courses but promises to be useful in future studies of assessment.

**Summary**

The use of RBIS in undergraduate life sciences education is increasing over time, but many students are still taught primarily through lecture and assessed using high-stakes exams focused on memorization. The evidence base available to determine what RBIS are being used and by whom is spotty and inadequate for drawing firm conclusions. It seems likely that many life sciences instructors are using some RBIS, but are not using these strategies extensively and might not be using them in a way that actually improves student learning (e.g., Andrews et al., 2011). There is wide agreement in life sciences education that student should be engaged in scientific practices, yet available
data indicates that such engagement may be even less common than RBIS like group work and peer instruction. CUREs are clearly gaining attention from life sciences educators, researchers, and policy makers, but data on the nature and extent of use of this approach is lacking. Finally, classroom assessment practices are even less well studied than classroom practices and undergraduate research experiences; they warrant additional investigation to determine whether assessments align with calls for reform in undergraduate life sciences education.

### REFORM IN INSTRUCTOR INTEREST, BELIEFS, AND ATTITUDES

Instructor interest, beliefs, and attitudes related to reform-based teaching have not been the exclusive focus of research on undergraduate life sciences education. The section below highlights examples of assessing instructor approaches to teaching and motivations within the context of instructor development and preparation for teaching.

#### Instructor Development and Preparation for Teaching

Several teaching professional development programs for undergraduate life sciences instructors operate on a national scale and have been assessed to determine program impacts on the use of RBIS. Outcomes assessments and studies of how these programs have evolved over time provide insight into how instructors can be supported to implement RBIS. The paragraphs below highlight the Summer Institute on Scientific Teaching and the Faculty Institute for Reformed Science Teaching. Other organizations also provide instructor development in the life sciences, including the National Center for Case Study Teaching in Science, and BioQUEST.

The Summer Institute (SI) on Scientific Teaching is a teaching professional development program for college science instructors. Originally this program focused only on biology instructors (Pfund et al., 2009). The basic goals and structure have persisted since its inception in 2004. College instructors attend a week-long institute focused on alignment between assessment and learning goals, diversity and inclusion, and the use of active-learning strategies (Handelsman, Miller & Pfund, 2007). Instructors disproportionately teach large introductory courses and attend in teams that include junior and senior faculty. During the institute, instructors work in teams to develop small instructional units called “teachable tidbits” and also participate in workshops that model RBIS and aim to build participants’ knowledge and skills relevant to RBIS. This program has expanded and evolved over time, and now includes regional institutes at various locations in the United States each summer, and mobile institutes that are institution-based. Over 2,200 instructors from over 350 institutions have participated in a Summer Institute, and 77% of these are life sciences faculty (Beth Louma, personal communication). Additionally, other programs have been modeled after the SI and designed to reach other instructor populations, including instructors at two-year institutions (e.g., Gregg, Ales, Pomarico, Wischusen, & Siebenaller, 2013).

The impact of participating in a SI on the use of RBIS continues to be investigated. Participants report learning gains and increased confidence in their ability to implement RBIS, and these gains persist two years after the SI (n = 68; Pfund et al., 2009). Participants also report using more active-learning and inclusive teaching practices two years after participating than at the start of the SI (Pfund et al., 2009). The observed
instructional practices of SI participants warrant additional study. A sample of 38 SI participants served as a comparison group in a study of the effectiveness of another teaching professional development program. Teaching practices were assessed using the Reformed Teaching Observation Protocol (RTOP) (Sawada et al., 2002). On average, SI participants’ teaching consisted of “lecture with some demonstration and minor student participation” and was not different from a sample of junior faculty who had participated in little or no teaching professional development (Ebert-May et al., 2015).

More recently, researchers compared instructional practices reported by students in courses taught by SI participants versus non-SI participants. Students in courses taught by SI participants reported a greater use of scientific teaching methods than did students in other biology classes (effect size as Cohen’s d = 0.71) (Durham, Knight & Couch, 2017). Given the large number of participants who have experienced the Summer Institute and the fact that the program is ongoing, there is great potential for future studies of how the use of RBIS among participants differs from that of non-participants.

The newest form of the Summer Institute is Mobile Summer Institutes (MoSIs), which are designed to address four categories of change strategies, including disseminating curriculum and pedagogy, developing reflective teachers, developing policy, and developing shared vision (Henderson, Beach, & Finkelstein, 2011). The MoSIs focus on a department and may more holistically foster change by involving a critical mass of instructors within a single department. MoSIs work with faculty and administrators to foster shared vision around scientific teaching and to set the stage for reflective teaching by building peer review into departmental practices. Assessment of program outcomes for MoSIs will be important.

Another teaching professional development program that aims to promote the use of RBIS among life sciences instructors is the Faculty Institutes for Reformed Science Teaching (FIRST). There have been multiple iterations of FIRST that have trained faculty and postdoctoral associates. FIRST II lasted three years and involved teams of faculty from the same department who met three-six times for a total of six-12 days of workshops (Ebert-May et al., 2011). As in the Summer Institutes, participants developed instructional materials that had clear learning goals and aligned assessments and that used active-learning strategies. The instructors who opted to participate in FIRST II also disproportionately taught introductory courses (Ebert-May et al., 2011). FIRST II participants reported that they were more knowledgeable and experienced with active-learning instruction after the program than before. They also reported using specific active-learning strategies regularly (Ebert-May et al., 2011). However, their observed instructional practices did not align with these optimistic reports. Most participants were observed to teach by lecture only or lecture with minor student participation (Ebert-May et al., 2011). Researchers further investigated what predicted these instructional practices and concluded that younger instructors were more likely to implement student-centered strategies.

A later iteration of the FIRST program, FIRST IV, focused on teaching professional development for postdoctoral fellows, and was more successful in promoting the use of RBIS. In addition to focusing on biology instructors earlier in their careers, FIRST IV focused on teaching practices rather than teacher tools and helped instructors to build an entire course rather than a teachable
Postdoctoral fellows participated in FIRST IV for two years, including two multiday summer workshops, teaching and revising a full course or part of a course, and receiving feedback on their teaching from mentors and peers. Participating in FIRST IV was associated with a reduction in instructor-centered perspectives on teaching, as measured by the Approaches to Teaching Inventory (Ebert-May et al., 2015). Additionally, participants reported being more knowledgeable and experienced with active-learning strategies after FIRST IV. Importantly, the observed instructional practices of FIRST IV were significantly more student-centered than the practices of FIRST II and SI participants, based on RTOP scores (Ebert-May et al., 2015). In another study of the program’s effectiveness, FIRST IV alumni were matched with another member of their department who had not participated in the program but was of the same rank and reported similar prior experience with teaching and active learning (Derting et al., 2016). The two groups (FIRST IV participants and matched non-FIRST IV participants) had similar scores on the Approaches to Teaching Inventory and similar perceptions of their teaching environment, including their department’s commitment to teaching. FIRST IV participants reported using active-learning strategies more frequently than did non-FIRST IV participants and they were observed to use more student-centered instructional practices. The differences in observed practices were large, Cohen’s $d = 1.64$, when comparing RTOP scores between FIRST IV and non-FIRST IV participants (Derting et al., 2016). About 85% of the assessment items used by FIRST IV participants assessed knowledge and comprehension in Bloom’s taxonomy, which is slightly better than national samples, but was no different from the matched comparison group (Ebert-May et al., 2015; Derting et al., 2016).

The directors of FIRST IV attribute the higher use of RBIS among participants to the fact that participants were early in their careers when they participated. They also highlight the benefits of designing an entire course within a mentored teaching professional development program and also to program elements that supported reflective teaching. Specifically, participants designed and taught a course after learning about RBIS, engaged in reflection with mentors and peers that included formative feedback on videos of their own instruction, and then revised and taught the course they designed a second time.

**Teaching Development for Graduate Students**

A research network called BioTAP (Biology Teaching Assistant Project) has brought additional attention to development and training for graduate teaching assistants in biology. The network aims to build research capacity for investigating teaching professional development for graduate teaching assistants (GTAs). One study from this group investigated institutional approaches for preparing biology GTAs at 71 institutions around the country (Schussler, Read, Marbach-Ad, Miller, & Ferzli, 2015). The vast majority of institutions required some mandatory teaching preparation for GTAs, but about half of the institutions required 10 or fewer hours of preparation (Schussler et al., 2015). This demonstrates increased levels of mandatory preparation since a 1997 survey showing that half of schools had no teaching preparation for biology GTAs (Rushin et al., 1997). Despite this improvement, fewer than 40% of institutions provided more intensive teaching development like a teaching certificate program, and only about 7% of institutions required 50 or more hours of teaching preparation. The most common
mandatory GTA preparation was a pre-semester course orientation and the most commonly emphasized topics included teaching policies, classroom management, course content, and teaching techniques (Schussler et al., 2015). Survey respondents were more satisfied with the teaching preparation that their institution provided to GTAs when it covered more topics, and often desired greater focus on pedagogical training and more observations of GTAs to provide formative feedback on teaching (Schussler et al., 2015).

**Department-level Teaching Development for Faculty**

Departmental initiatives to support teaching development may have some advantages over national workshops and programs. Departmental efforts are likely to have some level of administrative support that may lead to more instructors participating. Such efforts may also be better suited to blend multiple types of change strategies than outside teaching professional development. Specifically, departmental efforts may contribute to developing shared vision and policy in addition to developing reflective teachers and disseminating pedagogies (Henderson et al., 2011). This idea is corroborated by findings from a study of science faculty with educational specialties (SFES), which draws on survey and interview data with SFES showing that these faculty perceive that they have an influential role in initiating and shaping instructional change or other educational improvements in their own departments (Bush et al., 2008, 2011, 2013, 2016). Because college science faculty tend to have little preparation in teaching, the authors suggest that “the seeding of science departments with science education specialists holds promise for fostering change in science education from within” departments (Bush et al., 2016, Abstract).

One departmental effort to increase preparation for teaching among faculty explicitly used change strategies from each of the four categories (Henderson et al., 2011), and instructors adopted more RBIS. With funding from HHMI, the Department of Biology at San Francisco State University developed a department-wide, collaborative teaching professional development program called Biology Faculty Explorations in Scientific Teaching (Biology FEST). In addition to enhancing pedagogical expertise and supporting iterative change in biology instruction, Biology FEST aimed to engage faculty in regular discussions about student assessment evidence and build infrastructure for comprehensive teaching reform in the department (Owens et al., 2018). In brief, they started by holding two-hour workshops in which faculty could be exposed to Biology FEST with limited time commitment. Next, they held five-day FEST institutes that had many of the same goals as Summer Institutes, including a focus on scientific teaching, assessment, equity and diversity, and active learning. The Institutes also modeled active-learning strategies for participants throughout (Owens et al., 2018). Finally, Biology FEST included semester-long follow-up programs that brought together small groups of instructors to share personal insights about teaching but had some variable elements. In total, instructors who
attended a FEST institute and participated in a follow-up program spent nearly 100 hours working with colleagues about teaching (Owens et al., 2018).

Biology FEST developers carefully assessed the effectiveness of this program. The program was successful at reaching faculty: over 89% of instructors in the department, including tenure-track and lecturers, participated in a FEST institute, and over 80% of those who had participated in an institute also participated in a follow-up program. Multiple lines of evidence indicate that these participants increased their use of active-learning strategies, including their own reflective essays written after experiencing the institute and teaching a course, usage of a resource center that provides supplies often needed in an active-learning class, student reports of instructional practices, and classroom noise levels, as measured by Decibel Analysis for Research in Teaching (DART; Owens et al., 2017, Owens et al., 2018). Specifically, 22% of courses were using active-learning strategies in every class session and 81% were using active-learning strategies in at least half of their class sessions. These data cannot be easily compared to the RTOP data used to assess the Summer Institutes and FIRST programs, but offer a conservative estimate (Owens et al., 2018). Notably, instructors reported positive impacts on their teaching, confidence in their teaching, willingness to reflect on and make changes in their teaching, and willingness to take risks with their teaching. They also widely reported that their research was unaffected or positively affected by their participation in Biology FEST. Lastly, instructors felt a greater sense of belonging in the department and had experienced positive impacts on their relationships in the department. This was especially true for participants of follow-up programs (Owens et al., 2018). This outcome is notable because we would not expect it in other types of teaching professional development, but it may be important to sustained use of these strategies. Longitudinal evaluations of Biology FEST and programs like the Summer Institute and FIRST will help elucidate long-term impacts.

**Summary**

Prominent teaching professional development programs for life sciences instructors aim to promote a scientific approach to teaching, the use of active-learning strategies, and attention to diversity and inclusion. These programs have not always been rigorously assessed to determine their impact on the use of RBIS. When the teaching practices of participants are observed, results are mixed. Some programs seem to fall short of their goals because former participants still teach primarily by lecturing. One program produced instructors who taught using student-centered practices and attributed their success to a focus on early career instructors, designing and teaching a full course, and creating structure to foster critical reflection.
on teaching with mentors and peers. Another successful effort took place within a department and aimed to foster community and collaboration around teaching over extended periods of time. These findings provide some insight into important levers for change. Additional efforts to test the impact of these programs and specific components of the program will be important to advancing our understanding of how to support life sciences faculty in using RBIS. Given the observation that instructors earlier in their career may benefit disproportionately from teaching professional development, there will be particular value in better understanding teacher training for graduate students in the life sciences.

**DIVERSITY AND INCLUSION**

It is uniquely challenging to characterize the extent to which undergraduate life sciences education is attending to diversity and inclusion. I highlight a few indications that attention to diversity and inclusion may be increasing, including a special issue of *CBE—Life Sciences Education* that received a record number of submissions and one example of teacher development addressing diversity explicitly. However, my overall assessment is that much less attention has been paid directly to inclusion compared to more general efforts to foster change in classroom practices and teacher development.

A special issue of *CBE—Life Sciences Education* in September 2016 titled “Broadening Participation in the Life Sciences” published research and essays related to diversity and inclusion. The journal received over 120 submissions in response to their initial call for abstracts for this special issue, compared to a total of 194 submission in all of 2015 (Gibbs & Marsteller, 2016). This shows that the time was ripe for work examining diversity and that many efforts (and assessments of these efforts) were taking place and mature enough to be described in published research. Systematic investigations could help reveal the degree to which individual instructors, departments and universities, and national programs are attending to diversity and inclusion in undergraduate life sciences education.

There is some evidence that instructor development efforts are paying attention to diversity and inclusion. As described earlier, diversity is an explicit component of the Summer Institutes and FIRST programs, as well as the more localized Biology FEST initiative. However, the shape this takes has not always been well described or evaluated. Recently, the Summer Institute added a new focus on gender bias and assessed the effectiveness of their efforts. One 2-hour workshop specifically aimed to address gender bias among college faculty in the life sciences. The workshop actively engaged participants in thinking about social science literature regarding implicit bias, the benefits of inclusive teaching strategies, and shared responsibility for addressing diversity challenges (see details in Moss-Racusin et al., 2016). Researchers assessed participants’ awareness of diversity issues (n = 126), gender bias (n = 126), and readiness to take action (n = 78) two weeks prior to and two weeks after the workshop. Participants were more aware of gender diversity, exhibited lower gender bias, and were more ready to engage in behaviors addressing diversity challenges two weeks after the workshop (Moss-Rascusin et al., 2016). Explicitly training instructors to think differently about diversity will likely be important to fostering broader attention to issues of inclusion in undergraduate life sciences education. New research is needed to determine the extent to which teacher development programs are addressing inclusion and diversity head-on.
Summary
There are some indications that attention is being paid to diversity and inclusion in undergraduate life sciences education, but we lack evidence of the extent to which this is true. Additional research should be prioritized to determine the extent to which instructors, departments and institutions, and the discipline as a whole are attending to diversity and inclusion in undergraduate life sciences education.

FACTORS AND PROCESSES
The first section of this essay, which describes the extent to which research-based reform is occurring in undergraduate life sciences education, has previewed some potentially important levers for change. This section focuses explicitly on highlighting factors (i.e., levers) that impact the uptake of research-based reform. I address factors at three levels: the discipline as a whole, departments and institutions, and individual instructors. This multi-level look at levers is critical because decisions are made in each of these parts of the higher education system that ultimately impact what occurs in the classroom.

LEVERS FOR CHANGE IN LIFE SCIENCES AS A WHOLE
One far-reaching influence on life sciences education has been the publication of Vision & Change (AAAS, 2011). This document was produced as a result of conversations among over 500 instructors, administrators, representatives from professional societies, and students from around the country. It lays out a “vision” for reform in undergraduate life sciences education. The many meetings and collaborations that went into developing this document were key to its ultimate uptake because the life sciences are a vast discipline. Unlike most STEM disciplines, life sciences faculty are spread across many professional societies, which have different cultures and traditions. Additionally, life sciences instructors are often spread across many departments and even colleges. There is no national organization or meeting that brings together all, or even most, of the instructors who are ultimately responsible for leading undergraduate courses in the life sciences. Thus, many voices had to be heard and had to collaborate to reach a consensus document that would be seen as useful and relevant by instructors across specialties such as biochemistry, cellular biology, evolutionary biology, and ecosystem ecology.

Vision & Change calls for life sciences education to be grounded in five core concepts and six core competencies. For example, evolution and structure and function are two of the five core concepts, and an ability to apply the process of science and to communicate and collaborate with other disciplines are two of the six core competencies (AAAS, 2011). The report also envisions more widespread use of student-centered instruction, including backward design, active-learning instruction, frequent formative assessment, and hands-on undergraduate research experiences (AAAS, 2011). Beyond individual classrooms, Vision & Change recognizes a need for teaching professional development through professional societies and national programs for current and future college instructors, and Centers for Teaching and Learning on campuses (AAAS, 2011). It also highlights the current reward system in higher education as a barrier to achieving the proposed vision and calls for the development of additional models of how to reform reward systems (AAAS, 2011). The Vision & Change report is a guiding document that has served as a strong foundation on which many conversations and
initiatives can be built, from local efforts to develop learning objectives in a departmental degree program to national projects to promote reform.

There are multiple ways to document the impact of Vision & Change. I present two here. First, we can examine the degree to which Vision & Change is woven into grant proposals to fund initiatives and research related to reform in undergraduate life sciences. Program officers at the National Science Foundation analyzed biology-related proposals submitted to its Transforming Undergraduate Education in STEM (TUES) program (Vasaly, Feser, Lettrich, Correa, & Denniston, 2014). The percent of proposals citing Vision & Change rose from less than 1% in 2009 (when the final meeting to produce the report was held) to over 50% by 2013. They also searched words within proposals to determine if the proposed work integrated principles articulated within Vision & Change, including core concepts, core competencies, and student-centered instruction. These numbers also steadily increased, from 0% in 2009 to 22% in 2013 (Vasaly et al., 2014). We expect individuals submitting proposals to TUES to be more knowledgeable about research-based reform than a typical life sciences instructor, so these numbers do not represent awareness of Vision & Change across instructors. Nonetheless, these data indicate quick uptake of key ideas from Vision & Change by individuals planning reform initiatives and designing research that will eventually inform change efforts.

Another approach to examine the impact of Vision & Change is to look to ongoing initiatives that rely heavily on Vision & Change as a guiding framework. Here I highlight a few prominent, cross-institutional efforts, including the development of the BioCore Guide, BioMAPS, CourseSource, and PULSE. The core concepts described in Vision & Change have provided a basis for further articulating learning objectives and for designing assessments. For example, the BioCore Guide builds on the five core concepts by articulating overarching principles and specific statements that align with each core concept and also cross three main areas of the life sciences: cell, molecular, and developmental biology; physiology; and ecology and evolutionary biology (Brownell, Freeman, Wenderoth, & Crowe, 2014). This tool provides a much more detailed list of key concepts biology majors should learn, building on feedback from over 240 biologists and education researchers. Vision & Change (AAAS, 2011) provided an essential backbone on which to more specifically articulate key concepts through the BioCore tool. Important next steps include translating these key concepts, as well as the core competencies, into clearly articulated learning goals and objectives for undergraduate life sciences programs. This starts to fill a gap that might otherwise be addressed by a professional society. Vision & Change is an important starting place for this work because there is no single professional society that crosses all, or even most, of life sciences disciplines.

Another critical next step is being able to assess student knowledge of important learning objectives and goals with research-based tools. A cross-institutional research team is currently working to develop programmatic assessments for undergraduate life sciences programs that align with the key concepts in the BioCore Guide. This work has produced a capstone assessment for molecular biology degree programs (MBCA; Couch, Wood, & Knight, 2015) and a forthcoming capstone assessment for ecology and evolution, Evo/Eco-MAPS (Summers, personal
communication). The research group is also developing assessments addressing physiology, called Phys-MAPS, and genetics, called Gene-MAPS (J. Knight & B. Couch, personal communication). These instruments have been refined with data from over 13,000 students at over 60 institutions. The development of these resources would have been possible prior to Vision & Change, but the formal report helps researchers to argue for the relevance and need for such resources in undergraduate life sciences education and helps funding agencies prioritize such efforts.

Vision & Change proposed that reform in life sciences education would be facilitated by evidence-based teaching materials that were “collected in an ongoing and coordinated fashion and made easily accessible” (AAAS, 2011, pg. 53). CourseSource, an open-access journal of peer-reviewed teaching resources, answers this call. CourseSource publishes teaching materials that incorporate RBIS, focus on learning goals and objectives prioritized by various life sciences professional societies, and that include necessary descriptions and components for replicability and adaptation to other classrooms. As in the previous examples, Vision & Change facilitated the creation of CourseSource by articulating a need widely seen as important by the many contributors to Vision & Change.

PULSE (Partnership for Undergraduate Life Sciences Education) is another example of an initiative born as a result of Vision & Change. PULSE is a non-profit organization that was originally launched in 2012 by program officers from NSF, NIGMS/NIH, and HHMI. It aims to stimulate department-level implementation of Vision & Change across institution types. PULSE has developed tools to raise awareness and build capacity for reform within life sciences departments, tools to support planning and implementing reform, and tools for assessing reform and continued improvement. For example, PULSE created Vision & Change Rubrics to assess life sciences departments’ progress toward implementing Vision & Change in five areas: curriculum alignment, assessment, faculty practice/faculty support, infrastructure, and climate for change (Brancaccio-Taras et al., 2016). These rubrics can be useful for self-assessment and are the foundation for a recognition program currently being piloted in which departments peer-review each other’s progress toward implementing the principles of Vision & Change using the rubrics. PULSE continues to refine the rubrics to make them useful to the community. PULSE also has an ambassador program in which a trained PULSE ambassador visits a department to help facilitate conversations among administrators and life sciences instructors about implementing the recommendations of Vision & Change. These initiatives are promising, but rigorous evaluation data will be crucial to determining the success of PULSE.

Summary
At a national level, the most influential factor for research-based reform is the publication of Vision & Change (AAAS, 2011). Vision & Change was published less than a decade ago, and so the initiatives that followed from or were bolstered by its appearance are relatively new. As result, their full impact on the depth and breadth of reform in life sciences education has yet to be fully realized or rigorously evaluated. Austin (2018) summarizes the views of life science education leaders on these impacts and suggests issues for further study. What is clear is that Vision & Change has provided a vision of reform around which the life sciences community has coalesced. It provides legitimacy to efforts that predated the formal report and to newer efforts, especially those aimed at cutting content
in favor of core ideas and competencies and replacing instructor-centered teaching with a pervasive student-centered mentality. Outlining specific core concepts and competencies has provided a basic framework on which to build more clearly-articulated learning goals and objectives for undergraduate life sciences programs. It is unclear whether the impact of Vision & Change would have been the same in other STEM disciplines, but it is undeniable that this document has had a role in sparking, guiding, and legitimizing change in the life sciences.

**LEVERS FOR CHANGE IN ACADEMIC DEPARTMENTS AND INSTITUTIONS**

Departments are important sites of change in undergraduate life sciences education because instructors are typically hired, evaluated, and promoted primarily by departmental colleagues. Departments, even those within the discipline of life sciences, have distinct cultures (i.e., values, attitudes, practices and policies) that can impact undergraduate education and the uptake of research-based reform. For example, colleague-colleague interactions about undergraduate teaching varied across four life sciences departments within a single research institution (Andrews, Conaway, Zhao, & Dolan, 2016). Some departments had social networks about undergraduate teaching that were denser and more interconnected, meaning that more instructors talked to each other about teaching. The department that stood out most in these network characteristics was unique in at least two ways. First, the department head demonstrably valued undergraduate teaching. This leader ensured that both new and more senior instructors engaged in teaching development and mentoring, and publicly acknowledged innovative teaching at departmental events. Second, over half of the instructors in the department had participated in a week-long Summer Institute in their region. The department head played a significant role in encouraging faculty to attend. Additionally, the leader of the regional Summer Institute that they attended was an instructor in the department. Instructors reported in interviews that they thought their department valued effective, scientific teaching more than other life sciences departments at the institution, and that the instructors shared a “certain way of thinking” and “language” about teaching as a result of their common experiences in the Summer Institute (Andrews et al., 2016). This case highlights the department as a unit of potential and realized change, and also the critical role of department leadership.

Department chairs may be especially important players in life sciences education reform at larger institutions because they provide a key voice for the department and instructors in the department. At these institutions, multiple departments must be involved to substantially change undergraduate life sciences programs, because courses taught by a single department often serve majors across departments. For example, one initiative to reform life sciences education at a large research university involved nine distinct biology-related programs and departments (Matz & Jardeleza, 2016). Instructors at this institution struggled to make progress in reforming individual courses and working toward vertical alignment in the curriculum without a broad and shared vision of the purpose of biology education. This vision was eventually built as a result of many conversations among instructors that occurred within and across departments, as well as direction from policy enacted by administrators (i.e., bottom up and top down). Department chairs were critical
liaisons across departments and with administration to develop and shepherd this vision (Matz & Jardeleza, 2016). Additional research is necessary to fully understand the critical roles that department heads can play, and how these leaders may need to be trained and supported to be change agents for undergraduate education reform (e.g., Matz & Jardeleza, 2016).

Biology education researchers embedded within life sciences departments are also potentially impactful members of a department when it comes to research-based reform. Studies of such scholars across disciplines suggest that they engage substantially with efforts to improve teaching in their own departments (Bush et al., 2016). For example, the Biology FEST teaching professional development program relied on local science education expertise in the form of a biology education researcher and the department’s associate chair in charge of curriculum (Owen et al., 2018). These individuals had deep connections with other department members, giving them insight into the beliefs of individuals and the complex system in which they worked. They also had the expertise and experience with teaching and learning to design and lead long-term, intensive teaching professional development (Owens et al., 2018). Research at another institution found that biology education researchers in life sciences departments were seen by their colleagues as change agents for undergraduate education. Instructors reported changing their thinking and practices much more often as result of interacting with biology education researchers than from interacting with other departmental colleagues (Andrews et al., 2016). Each of these studies is essentially a case study of one or a few departments. Investigations on a much larger scale are necessary to better understand the role that biology education researchers can have in research-based reform.

In addition to individuals, structures may be important levers for change within departments and institutions. One of the most commonly discussed structures is the reward system. Which rewards are important for promoting and sustaining research-based change is not clear. Tenure, promotion, and salary raises are formal and far-reaching rewards. Praise and respect from colleagues are more informal, but potentially highly impactful. Rewards may also include monetary compensation for time devoted to teaching professional development or accrued as a result of teaching awards. Beyond individual instructors, funds may be available for departments or programs that meet specific objectives (e.g., Matz & Jardeleza, 2016). Though rewards are commonly thought to be important barriers, and potential levers for change, we lack research to clearly illuminate what rewards are important to whom and in what contexts. Indeed, there are indications that we should carefully test our assumptions about the impact of rewards. For example, instructors were generally motivated to participate in Biology FEST due to their own teaching and their interest in building community, rather than monetary compensation (Owens et al., 2018). Money and promotion were the most commonly named incentives that biology instructors reported would entice them to seek feedback on their teaching, followed closely by improved student learning and not needing any incentive (Brickman, Gormally, & Martella, 2016). Clearly, there is an urgent need for more research within life sciences and across STEM to investigate the role of the reward system in research-based reform. This work is particularly pressing because it will likely require long-term studies and studies across contexts, and because
of the important role a reward system is hypothesized to play.

Other departmental structures may also be important to the uptake of RBIS, including opportunities colleagues have to collaborate in their teaching, hiring practices, communication of the expectations for teaching, and mechanisms for providing formative feedback for teaching (e.g., Brickman et al., 2016; Dennin et al., 2017). The impact of these factors on teaching practices has yet to be investigated but existing work suggests these structures are not designed to facilitate research-based reform. A national survey of 400 college biology instructors found that instructors valued peer evaluations of their teaching, but less than half of instructors at doctoral-granting institutions participated in this type of teaching evaluation (Brickman et al., 2016). Additionally, it was uncommon for biology instructors to use teaching portfolios or student performance data to inform their teaching (Brickman et al., 2016). There is substantial room for research to understand the departmental practices and policies that influence research-based reform.

Summary
There are many potential levers for change at the level of the department, college, and institution that warrant further investigation. Most studies examine a single department or institution, so making generalizations requires many cases to be richly described and then compared. A small body of work indicates that department leadership plays an important role in setting a standard for how undergraduate teaching and teacher development is valued. Departmental leadership also provides a voice for instructors across departments and with upper administration. This role of leadership may be especially important in the life sciences, where instructors are often spread across departments and even colleges. Biology education researchers and other education specialists may also be a lever for change within a department. Such an impact likely depends on the individual and context, making larger scale studies important. Lastly, the reward system and other departmental policies and practices are widely thought to be important, but these hypotheses remain mostly uninvestigated in life sciences education.

LEVERS FOR CHANGE WITH INDIVIDUAL INSTRUCTORS
Ultimately, we hope to see research-based reform in classrooms. Therefore, individual instructors are at the front lines of reform. Research that examines factors affecting reform among instructors provides insights, but many important questions remain. Owens et al. (2018) argue that we are often guided, implicitly or explicitly, by a faculty deficit model that assumes that the reasons instructors do not adopt new teaching practices lies primarily within the individuals, rather than problems in the higher education system or society as a whole. Yet their observations of faculty participation and growth within the FEST program did not support a faculty deficit model. Instructors were motivated to participate in this program, and the majority of the Department of Biology participated in over 100 hours of teaching development (Owens et al., 2018). With that in mind, I reviewed relevant research that examines factors internal and external to the instructor that may influence the adoption and sustained, effective use of RBIS.

One important lever for change among individual instructors is their own perceptions of what is happening in their classroom. Importantly, life sciences instructors adopting RBIS may be more motivated by
personal experiences than by empirical evidence. Adopters of case study teaching often felt dissatisfied with teaching primarily via lectures and this motivated them to seek alternative approaches (e.g., Andrews & Lemons, 2015). Though they were generally aware of empirical evidence demonstrating the effectiveness of RBIS, it was not that evidence that convinced them to change their teaching. This has also been observed in other STEM disciplines and highlights the important role that personal perceptions dissatisfactions can have in reform (Gess-Newsome, Southerland, Johnston, & Woodbury, 2003; Marbach-Ad & Rietschel, 2016). Instructors also rely on their own observations of student interest and engagement in class and on feedback volunteered from students to judge their own effectiveness (e.g., Andrews & Lemons, 2015; Marbach-Ad & Rietschel, 2016). Personal experiences and feedback from students contribute not just to initial adoption, but also to iterative revisions of the implementation of RBIS (e.g., Marbach-Ad & Rietschel, 2016).

Adopting RBIS is a long-term process and that life sciences instructors make small improvements semester after semester (e.g., Andrews & Lemons, 2016, Marbach-Ad & Rietschel, 2016). In fact, it is likely unreasonable to assume that new users of RBIS will be successful during their first implementations of these strategies. A willingness to try and fail, access to knowledgeable others, and colleagues with whom instructors collaborate and innovate in their teaching may all be important conditions for sustained reform (e.g., Andrews & Lemons, 2016; Marbach-Ad & Rietschel 2016).

Overall, perceived affordances may be more important to individual instructors’ actual implementation of RBIS than barriers. Researchers surveyed 485 former Summer Institute participants about the affordances and barriers they perceived to using RBIS. They also asked them to report their use of practices taught in the Summer Institute (e.g., active learning, assessment approaches, inclusive teaching practices). The affordances that instructors perceived were strongly and positively correlated with their use of RBIS practices and the barriers they perceived were weakly negatively correlated with these practices (Bathgate et al., in review). Additionally, the affordances that instructors perceived were highly interrelated, so if instructors experienced one, they experienced many others. In contrast, barriers were not as interrelated, so removing one barrier might have little effect on other barriers perceived. This work does not identify the affordances that were most impactful, but the most common related to deriving personal pleasure and satisfaction from using RBIS, appreciating the chance to get to know students, supportive colleagues, and students engaging with each other during in-class activities (Bathgate et al., in review).

One proposed barrier to the adoption of RBIS that has received a lot of attention in the biology education research community is the professional identity of scientists. Professional identity includes how a person defines themselves professionally, including workplace values, roles, and responsibilities, and how the person is seen by those around them (Hall & Burns, 2009). Brownell and Tanner (2012) proposed three tensions between a professional identity as a scientist and the adoption of RBIS. The first tension is that training as a scientist promotes the development of a professional identity as a researcher, but not as a college teacher (Brownell & Tanner, 2012). The second tension is that scientists may be afraid to “come out” as teachers because they fear they will not be taken seriously by the larger scientific community (Brownell & Tanner, 2012). The
third tension is that the professional culture of science considers teaching to be lower status than research, and instructors may feel that to be seen as “real” scientists they need to shy away from spending time on their teaching and teaching development (e.g. Thiry, Laursen, & Liston, 2007; Brownell & Tanner, 2012). The essay proposing these ideas is one of the most commonly-cited papers published in CBE—Life Sciences Education and has been cited over 200 times. This is evidence that the ideas resonate with the community and may be aligned with researchers’ personal experiences and perceptions. How professional identity develops during and after graduate school and the impact this has on the adoption of RBIS has yet to be empirically investigated but should be a priority.

Research investigating the experiences of graduate students in the sciences indicates that students perceive these tensions in their training programs and have to be resilient to pursue teaching interests in this environment (e.g., Thiry et al., 2007; Connolly, 2010; Lane, Hardison, Simon & Andrews, 2018). Many unanswered questions remain, however. The role of professional identity is likely to be complex because it lies at the intersection of the culture of a discipline, the culture of departments and institutions, and an individual’s self-perceptions.

One notable gap in research on levers for change for life sciences instructors are investigations of instructors’ knowledge and skills for teaching. The knowledge an instructor employs influences how he or she implements RBIS and implementation influences student outcomes (e.g., Park et al., 2011; Santagata & Yeh, 2014). Planning and implementing RBIS likely requires knowledge well beyond content knowledge, including generalized knowledge of teaching and learning (pedagogical knowledge) and topic-specific knowledge of teaching and learning (pedagogical content knowledge). Furthermore, using RBIS requires different knowledge than lecturing. While a lecturer can prepare an entire lesson before class and enact it without much deviation, an instructor using RBIS engages students in examining and articulating their own thinking during class (Wagner, Speer & Rossa. 2007). An investigation of knowledge used by life sciences instructors who report employing active-learning strategies in large courses found significant differences between highly effective (i.e., expert) instructors and instructors new to RBIS (i.e., novices). Instructors analyzed active-learning lessons in large biology courses, and researchers compared what they noticed in their analyses and how they reasoned about what they noticed (Auerbach, Higgins, Brickman, & Andrews, 2018). Experts and novices differed in what they noticed, with experts more commonly considering how to hold students accountable, topic-specific student difficulties, eliciting and responding to student thinking, and opportunities students have to generate their own ideas and work. Experts were also better able to support their lesson analyses with reasoning. This research, work in other STEM disciplines, and a large body of research studying K12 instructors indicates that teacher knowledge should not be overlooked as we consider how to support instructors in adopting and effectively implementing RBIS.

Lastly, investigations of teacher development programs highlight potential levers. Early-career instructors may be more likely to adopt RBIS (e.g., Ebert-May et al., 2015). If so, we may see the largest long-term impact from focusing on the development of graduate students and postdocs. Additionally, programs that involve instructors in teaching training over semesters or even years,
and that aim to inform the development or revision of entire courses seem to be impactful than intensive, but short programs (e.g., Ebert-May et al., 2015; Owens et al., 2018). This long-term engagement can be achieved from programs within departments. Additional benefits of programs embedded within departments and institutions include building community around teaching and slowly shifting cultural values and expectations regarding teaching (e.g., Owens et al., 2018). Additionally, teaching preparation for early career instructors may be best provided within the institution or department so that long-term support is practical. This raises questions about the expertise in undergraduate life sciences education needed within departments and institutions and highlights a potentially important role for education specialists.

**Summary**

Many potential levers for change at the level of individual instructors warrant further investigations. Here I have proposed levers that prior studies indicate may be important. Instructors seem to be motivated primarily by their own perceptions of their teaching, and especially by dissatisfactions. Further research can aim to better understand approaches for fostering dissatisfaction among life sciences instructors. Dissatisfaction alone, of course, is not enough. Instructors also need to be able to prioritize changing their teaching. This may be impacted by the priorities they perceive in their departmental and disciplinary culture, and also by the opportunities they have to collaborate with colleagues around teaching. Because instructors rely on their own perceptions of their effectiveness and the responses of their students to make decisions, it may be important that they have sufficient knowledge, skills, and resources to achieve at least some success with RBIS early in adoption. Achieving sustained and highly effective implementation of RBIS likely requires more highly developed knowledge. Much research has been dedicated to studying undergraduates as learners of biology, including their motivations, cognitive processes, and approaches for facilitating their learning. We also need research to understand life sciences instructors as learners of RBIS.

**CONCLUSION**

The extent of research-based reform in undergraduate life sciences education has increased over the last decade or two, but still falls short of being widespread or revolutionary. Strategies like polling students and engaging them in class discussions are being used more commonly than strategies that engage students in the practices of sciences. CUREs are an attractive approach for solving this problem but are not yet used widely enough to impact most undergraduates. Reform in how we assess students has not kept up with changes in how students are taught, potentially substantially undermining attempts to change students’ opportunities for learning. New tools, like 3D-LAP, can help us recognize and solve this problem. We are learning from evaluations of teaching preparation programs, but it is unclear how many life sciences instructors are actually engaging in substantial teacher training and the impact of such involvement. Effective examples of teaching professional development should guide additional efforts. Finally, despite persistent and focused attention from funding agencies on broadening participation in STEM, is it unclear that typical life sciences instructors are paying much attention to inclusion and diversity. This is another important role that teaching preparation for current and future instructors can play.
The factors and processes that influence research-based reform in the life sciences should be considered an area in need of immediate and long-term attention by researchers. The current research base is sparse and answering important questions requires many localized studies, intensive national-scale studies, and longitudinal work. Existing research indicates that Vision & Change has had an impact, and that structures exist that can increase its impact over time. Departments will also be central to change, especially leaderships' advocacy for valuing teaching and teaching preparation. Other practices and policies within departments and institutions are likely important, but this has largely gone uninvestigated in undergraduate life sciences reform. Factors at the level of the individual instructor have received more attention and indicate that we should think more about what facilitates change than what prevents change. Personal perceptions of instructors weigh heavily in decision-making and should be considered an important lever that can be exploited to foster change. The role of professional identity has piqued interest but needs empirical attention. Evidence indicates teacher knowledge will be important to effective use of RBIS, but many questions are unanswered. Individuals seem to benefit from teacher preparation that occurs early in their career and is long-term and closely linked to their actual teaching responsibilities.

As in any exciting area of scholarship and important societal issue, there are many more unanswered questions than there are firm findings when it comes to understanding the extent of research-based reform in undergraduate life sciences education and the levers that impact this reform. I hope this report will act as an invitation to both established and new scholars to ask and answer questions to fill in the many gaps that exist in what we currently know in this area.

REFERENCES


Websites that provided information included in this report:

Biology Teaching Assistant Program, BioTAP, https://biotap.utk.edu/

CourseSource, https://www.coursesource.org/

CUREnet, https://curenet.cns.utexas.edu/content/about-curenet

Partnership for Undergraduate Life Sciences Education, https://www.pulse-community.org/

Summer Institutes on Scientific Teaching, https://www.summerinstitutes.org/
LEVERS FOR CHANGE: An assessment of progress on changing STEM instruction

Marilyne Stains, University of Nebraska-Lincoln

With valuable input from: Cheryl Bailey, Renée Cole, Ashley Donovan, Andrew Feig, Tom Greenbowe, Stacey Lowery-Bretz, Marcy Dubroff, Jon Holmes, Britton Miller, Rick Moog, Emily Moore, Erika Offerdhal, Jeff Raker, Susan Rowland, Susan Shadle, Joanne Stewart, Terri Taylor, Sonia Underwood, Jodi Wesemann

INTRODUCTION

Instructional reforms of the undergraduate chemistry and biochemistry curricula have been ongoing for decades. New content pieces, assessment tools, and instructional strategies have been developed based on empirical understanding of how students learn in these disciplines. Although not grounded in empirical data, Gabel in 1999 offered a pessimistic assessment of changes that had occurred in the 20th century for the chemistry curriculum and called for more action (Gabel, 1999):

Unfortunately, chemistry education research in the 20th century has had little influence on the way chemistry is taught. The changes that have occurred in textbooks during the past four decades have not been driven to any great extent by research findings. Although chemistry education researchers have identified common misconceptions for almost every topic taught in introductory science courses, probably nine out of ten instructors are not aware of these misconceptions or do not utilize ways to counteract them in instruction. [...] There is a need to focus on ways to incorporate current research findings into the teaching and learning of chemistry today (pp. 552, 553).

This report addresses the extent to which instructional changes have propagated over
the last two decades into the chemistry and biochemistry undergraduate curricula. A critical look at the available evidence to assess current state of reform will also be discussed, along with factors that influence propagation. The questions that the report was asked to address follows:

1. To what extent, and in what ways, is research-based reform in undergraduate STEM instruction occurring in this discipline? How do you know?
   ▲ What evidence from research and practice is available about the nature and extent of implementation of research-based reforms in STEM instruction? What evidence is missing or inadequate?
   ▲ What does that evidence tell us?
   ▲ What changes have been observed or reported, in what domains? Which types of change seem to be more established, and which less so?

2. What factors and processes have influenced the depth, breadth and impact of these changes?
   ▲ What factors—events, documents, people, movements, contextual circumstances—have influenced, in positive or negative ways, the depth, breadth, and impact of these changes?
   ▲ What lessons can be learned from the available evidence about ways to measure or monitor the nature and extent of research-based reform in undergraduate STEM instruction?
   ▲ What do we NOT know that is important about the extent of implementation of research-based reform in STEM instruction or about factors affecting such changes?

To answer these questions, we consulted peer-reviewed journals in these disciplines (Journal of Chemical Education and Biochemistry and Molecular Biology Education) focusing on publications over the last twenty years, websites of the disciplinary professional societies (American Chemical Society and American Society for Biochemistry and Molecular Biology), reports published by the National Academies Press, developers of research-based instructional strategies (RBIS) and leaders of instructional reforms in these disciplines, ACS Symposium book series, and websites associated with specific RBIS. We also conducted literature searches on Google Scholar. We specifically focused on studies conducted on a large scale. We included instructional practices such as RBIS, but also identified research-based curricula. We attempted to be as comprehensive as possible but recognized that we may have missed information.

**RESEARCH-BASED REFORMS**

This first section introduces the various ways research-based instructional strategies have been introduced in classroom practices, professional development programs, assessment, and available resources, as well as how professional societies and others have evaluated and/or promoted improved chemistry and biochemistry teaching and learning.

**CLASSROOM PRACTICES**

While several large-scale studies conducted within the last decade provide insight into the instruction students experience in the lecture portion of chemistry or biochemistry courses, the Higher Education Research Institute (HERI) is the only organization since 2004 to conduct national surveys of full-time faculty teaching at the undergraduate level.
Part of the survey focuses on instructional practices, classified by broad disciplines. Chemistry is included under the physical sciences umbrella that includes astronomy, atmospheric sciences, geological and earth sciences, and physics. Biochemistry is included in the biological sciences that includes biology, botany, genetics, microbiology, physiology, and zoology. Although these data are not specific to chemistry and biochemistry, they provide the only national and longitudinal outlook on instructional practices that we could find.

Eagan analyzed the HERI surveys from 2004 to 2014 to focus specifically on changes over time in aggregate faculty self-report on their teaching practices (Eagan, 2016). His analysis showed:

▲ Similarly small decline in physical and biological sciences faculty reports of using extensive lecturing in all or most of their courses from slightly over 80% in 2004 to about 65% in 2014;

▲ No change in use of class discussion in all or most courses for physical sciences (about 60% throughout the 10 years) and a small increase for biological sciences especially since 2010 (from 65% to 70% throughout the 10 years);

▲ A sharp increase for both types of faculty in the use of electronic quizzes with immediate feedback in all or most courses (e.g., clicker questions) from 10% in 2004 to 20-30% in 2014, as well as for the use of student inquiry to drive learning in all or most courses especially since 2010 (from about 35% in 2004 to almost 50% in 2014 in the physical sciences and 38% to 56% in the biological sciences);

▲ And the use of group projects in all or most of courses especially since 2010 (from 20% in 2004 to about 33% in 2014 for physical sciences and from 29% to 51% for the biological sciences).

Overall these results indicate a moderate shift toward student-centered practices over that decade, with a greater growth rate developing in 2010.

In contrast to the HERI study, which provides a large-grained understanding of chemistry and biochemistry instructional practices, two studies have looked at the instructional practices implemented in specific chemistry courses. First, a national study explored the instructional practices of physical chemistry instructors from chemistry programs certified by the American Chemical Society (ACS) (Fox, 2016; Fox & Roehrig, 2015). Surveys were collected in spring 2014 from 331 instructors from a range of institution types. Instructors were asked
to identify their classroom environment as instructor-centered, student-centered, or both, where definitions of these terms were provided. The majority of respondents (79%) reported providing an instructor-centered environment, a minority (19%) reported using an equal mix of instructor- and student-centered environments, and only 2% reported providing a student-centered environment. An analysis of type of environment versus institution type revealed that the instructors providing the student-centered environment were mostly from baccalaureate institutions (80%) and master’s institutions (20%). On the other end, 41% of the instructors providing the instructor-centered environment were from doctorate-granting institutions, 36% from master’s and 23% from baccalaureate institutions. The second study also involved a national survey but focused on instructors teaching inorganic chemistry courses (Reisner et al., 2015). Over 400 people responded to the online survey and represented a variety of institution types, although doctorate-granting institutions were overrepresented. Instructors were from ACS and non-ACS approved programs. The majority of instructors (72%) reported using in-class activities along with lecture; only a minority relied solely on lecture (20%). Lecture-only practices were more prominent in doctorate-granting institutions.

Another way to provide a fine-grained look at instructional practices is to focus on specific types of instructors. Stains, as part of her work evaluating the Cottrell Collaborative New Faculty Workshop (Baker et al., 2014), invited 555 assistant professors from doctorate-granting institutions to take an online survey probing their knowledge and use of RBIS (Pilarz, Chakraverty, & Stains, 2014). A little over a fifth (126) responded. Almost half (44%) were faculty who were starting their academic career, 35% had just finished their first year, and 21% were anywhere between their second and fifth year as assistant professor. Respondents were asked to indicate their level of awareness and use of eighteen RBIS. Starting and first-year faculty had knowledge of, on average, half of the RBIS, while those in their 2nd-5th years knew on average 11. The adoption rate was extremely small: on average, first-year faculty had implemented one RBIS and more experienced faculty had implemented 2.2 RBIS. The sample surveyed in this study is likely not representative of the population of assistant professors at doctorate-granting institutions and likely overrepresents those with a higher interest in teaching than average. Therefore, these findings present the best-case scenario and adoption is likely lower among the overall population.

Other studies that have attempted to capture instructional practices of chemistry and biochemistry instructors on a large scale are based on classroom observations. Analyses in these studies, which focused mostly on STEM instructors at doctorate-granting institutions, are based on videos of one or more lectures collected from each instructor participating in the study. In one set of studies, videos are analyzed.
using the Classroom Observation Protocol for Undergraduate STEM (COPUS) (Smith, Jones, Gilbert, & Wieman, 2013). COPUS requires the observer to mark whether one or more of 12 instructor behaviors and 13 student behaviors occur during each two-minute interval of a class. Authors of these studies relied on a sub-sample of the COPUS behaviors to run cluster analyses and identify broad instructional styles (Lund et al., 2015; Stains et al., 2018). We focus on the results of the latest study because it provides a better snapshot of the instructional landscape in STEM: 2,008 observations from over 500 STEM instructors including 121 in chemistry, compared to 269 observations from 71 STEM instructors including 39 in chemistry. Only 10 observations from three biochemistry instructors were part of this sample and, thus, results associated with biochemistry are not discussed here.

The cluster analysis led to the identification of seven instructional profiles, which could be further classified into three broad instructional styles (Stains, et al., 2018): Didactic (80% or more of class time is devoted to lecturing), Interactive Lecture (lecture supplemented with student-centered strategies), and Student-Centered (a large portion of class is spent on student-centered strategies). The authors investigated the distribution of these three broad styles across all seven STEM disciplines represented in the study. They found that 63.5% of the chemistry observations were classified as Didactic, 25.4% as Interactive Lecture, and 11.1% as Student-Centered. The chemistry observations had the highest representation of the Didactic style and the lowest representation of the Student-Centered style when compared to the other disciplines. The results of this study are similar to those from smaller-scale studies that used COPUS (Lund, et al., 2015) and the Teaching Dimensions Observation Protocol, the antecedent to COPUS (Hora & Ferrare, 2013; 2014).

Several studies have also explored the adoption of classroom response systems (CRS) such as clickers (Emenike & Holme, 2012; Gibbons et al., 2017). CRS have been demonstrated to enhance student learning when paired with student-centered pedagogies such as Peer Instruction (Crouch & Mazur, 2001; Vickrey, Rosploch, Rahmanian, Pilarz, & Stains, 2015). Emenike and Holme (2012) conducted a national survey on assessment in 2010. Data were collected from over 1,500 chemistry instructors from two-year, four-year, and doctoral-granting institutions. They found that 18.6% of these instructors were using CRS in their courses. Instructors at doctoral-granting institutions were more likely to use clickers compared to those at two-year and four-year institutions. Use of CRS was highest among chemistry education faculty (25.9%) and lowest among organic and biochemistry faculty (16.5% for both). Gibbons et al. (2017) revisited the original study and investigated adoption of CRS in various contexts. They implemented a stratified random sampling method across institution types (i.e., public versus private, highest degree granted in chemistry), which led them to collect surveys from over 1,200 instructors in 2016. Their results indicate that 21.1% of the surveyed instructors used CRS, most of them using it every class meeting. Therefore, minimal uptake of CRS has taken place over the six-year period separating these two studies. Gibbons et al. also found that CRS adoption is higher at public institutions, in large classes, and introductory courses. These studies also do not provide information as to the pedagogical approaches used in tandem with the CRS.
PROFESSIONAL SOCIETIES

Reform uptake in classroom practices and assessment at the discipline level was also explored by analyzing documents from the professional societies for chemists, the American Chemical Society (ACS), and for biochemists, the American Society for Biochemistry and Molecular Biology (ASBMB).

American Chemical Society
The ACS has had guidelines for the certification of bachelor’s degree programs for decades. These guidelines are revised every few years. Over the last two decades, four revisions have been made (ACS, 1999, 2003, 2008, 2015). We focused our analysis on changes in the guidelines related to instruction and assessment. Guidelines regarding instructional methods evolved drastically over time, with a major shift occurring in 2008. In 1999, the committee compiling the guidelines devoted only one section to pedagogy, briefly mentioning some RBIS, and encouraging the use of a wide range of practices:

Alternative Pedagogies. Much experimentation occurs in the teaching of chemistry as well as in the organization of the content and the formulation of laboratory experiments. This experimentation involves, for example, laboratory-driven instruction, problem-solving formats, and group learning. The Committee considers the guidelines to be consistent with using a wide range of pedagogies (ACS, 1999, p. 13).

In 2003, the guidelines acknowledged the growing research on how students learn, provided more examples of RBIS, and promoted the use of innovative pedagogies:

Curricular Innovation and Pedagogical Approaches. The Committee considers the guidelines to be consistent with using a wide range of pedagogies. Research in teaching and learning is generating an increasingly detailed picture of how students learn in their courses. Chemistry faculty are using this knowledge to improve student learning by, for example, having students build from their past experiences, using laboratory experiences to drive course instruction, organizing course content around topics of particular relevance to students, having students work in groups to build knowledge, and having students communicate their learning and results to others. The Committee encourages innovative pedagogical efforts and curriculum development (ACS, 2003, p. 14).

In 2008, the pedagogical statement was moved from the “Commentary on Curriculum Requirements” section to the “Curriculum” section and it was the first sub-section. Importantly the language shifted drastically with an expectation to use RBIS, a requirement to provide instructors with pedagogical professional development, and a call for programs to self-assess the effectiveness of the pedagogical strategies employed in the department:

5.1 Pedagogy. An approved program should use effective pedagogy in classroom and laboratory course work. Programs should teach their courses in a challenging, engaging, and inclusive manner that accommodates a variety of learning styles. Additionally, a program should provide opportunities for faculty to maintain their knowledge of best practices in chemistry education and modern theories of learning and cognition in science. An approved program should regularly review its pedagogical approaches to ensure that it provides excellent content and builds skills that students need to be effective professionals. Faculty should incorporate pedagogies that have been shown to be effective in undergraduate chemistry education. Examples include problem- or
inquiry-based learning, peer-led instruction, group learning, learning communities or networks, writing throughout the curriculum, and technology-aided instruction. Laboratory work provides a particularly attractive opportunity for inquiry-driven and open-ended investigations that promote independent thinking, critical thinking and reasoning, and a perspective of chemistry as a scientific process of discovery (ACS, 2008, p. 8 & 9).

The 2008 guidelines also added a new section focused on the skills students should gain through the program (e.g., problem solving, use of chemical literature, teamwork) and a requirement for programs to assess students’ development of these skills. The 2015 guidelines made only minor changes to both of these sections (ACS, 2015).

The ACS surveyed all ACS-approved programs in 2012 to capture the impact of the 2008 guidelines on their practices (American Chemical Society, 2013). They received surveys from 427 programs. Only a minority (12%) indicated employing “new pedagogical approaches” as a result of the new guidelines.

American Society for Biochemistry and Molecular Biology

ASBMB has been intentionally focused on transforming instructional practices in biochemistry for over a decade (Mattos et al., 2013). In 2008, the ASBMB Education and Professional Development Committee started the process of developing an accreditation program for bachelor’s degrees in biochemistry and molecular biology (Kennelly & Bell, 2011). Although there was some skepticism (Linn, 2011), the program was adopted and first implemented in 2013. The program was developed, tested, and refined by members of the biochemistry and molecular biology educational community. The accreditation document explicitly states that they see this accreditation process as “a powerful vehicle by which the ASBMB can actively and visibly promote excellence and innovation in undergraduate BMB education” (ASBMB, 2017d). Key characteristics of this accreditation program are curricular recommendations based on core concepts and learning objectives (ASBMB, 2017a,c), rather than a list of required courses, and nationally recognized assessment tools to measure students’ performance against these learning objectives (ASBMB, 2017b) and to assess program effectiveness (ASBMB, 2017a). The accreditation document identifies active learning as a key characteristic of the curriculum of accredited programs (ASBMB, 2017a):

Students also should have some level of active learning in one or more of its many forms. This approach to teaching can engage students in either the classroom or the teaching laboratory. There are many approaches of active learning that provide opportunities for students to think critically and that promote student engagement, synthesis, analysis, team skills and problem solving. The ASBMB has a long history of workshops and sessions providing support for those interested in developing active learning strategies.

As of late 2017, 68 institutions of all types had been accredited across 29 states (Kennelly, 2017).

Summary

Chemistry instructors rely heavily on didactic strategies, particularly in doctorate-granting institutions. CRS that can potentially promote student-centered practices have not had a large uptake overall, and the spread is uneven across institution types. There is little evidence about the teaching practices of biochemistry instructors. Professional societies have moved toward requiring the implementation of RBIS in their certification or accreditation programs, however, the impact of these changes is not clear.
Evidence of instructional practices in chemistry and biochemistry is limited in type (either self-reported surveys or classroom observations) and scope. For example, most studies do not have a representative sample of instructors across institution types and offer limited understanding of teaching practices in various sub-disciplines of chemistry and across the undergraduate curriculum. The available studies report only on the classroom/lecture portion of the course. Consequently, little is known about student experiences in other components of a course such as recitation, online environments associated with the course, and the laboratory. Moreover, the evidence gathered to date is mostly behavioral; the quality of the instruction has not been captured. Finally, evidence regarding the impact of changes in the ACS requirements to implement active learning in their certified program is limited, but it seems that evidence is being collected in biochemistry.

CLASSROOM PRACTICES IN THE LABORATORY

Instructional practices carried out in the laboratory have not been studied extensively over the last two decades. One study explored the instruction provided in first-year general chemistry laboratory courses (Hilosky, Sutman, & Schmuckler, 1998). They observed 24 instructors from 16 institutions of higher education. Analyses were carried out using a modified version of the Science Teacher Behavior Inventory (Vickery, 1968). They found that lab instructors spent most of their time “supervising students’ laboratory work” and little time engaging students in explanations of concepts involved in experiments. They concluded that the instructional strategies most frequently implemented would not promote higher-order thinking processes. Interestingly, findings were similar in the major and non-major laboratory courses.
A second study investigated STEM Teaching Assistants’ (TAs) educational and instructional environment at a doctorate-granting institution in which training programs were provided to TAs (Luft, Kurdziel, Roehrig, & Turner, 2004). Five chemistry TAs were among the study participants. Through interviews, observations and focus groups, the authors also found a dominance of didactic practices and few opportunities for students to develop scientific thinking.

A third study looked into the practices of 14 new chemistry TAs in a doctorate-granting institution by surveying undergraduates in these laboratories at the end of the semester and by having faculty analyze videotapes of the TAs’ lab sessions (Rodrigues & Bond-Robinson, 2006). Data indicated discrepancies of assessment of the learning environment between undergraduate students and faculty, with undergraduate students being unable to assess TAs’ integration of chemical concepts. Faculty observations indicated that most TAs excelled at managing the laboratory but failed to promote conceptual understanding.

More recently, Velasco et al. (2016) employed a modified version of COPUS—the Laboratory Observation Protocol for Undergraduate STEM (LOPUS)—to characterize the instructional practices enacted in 19 general chemistry laboratory sessions taught by 15 TAs at a doctorate-granting institution. They found that TAs spent most of their time talking with individual groups or students (75% of the 2-min time intervals on average); other behaviors such as monitoring students and asking questions to individual students or groups were less represented (less than 30% of 2-min time intervals) and this representation varied widely from session to session. A hierarchical cluster analysis led to the identification of four instructional styles that were equally distributed across the 19 observations: the ‘waiters’ were TAs who wait for students to ask them questions; the ‘busy bees’ went from lab station to lab station to assist students; the ‘observers’ were focused on monitoring students; and the ‘guides-on-the-side’ questioned and praised students.

The extent to which inquiry teaching is being implemented in STEM laboratories was investigated through an analysis of first-year laboratory exercises published in STEM practitioner-oriented peer-reviewed journals from 2000 through 2010 (DeChenne, Carew, & Stains, 2014). They identified the authors of laboratory exercises published in the Journal of Chemical Education and found that 87% were from two main groups, faculty with primary teaching positions or bench chemists. The researchers argued that the laboratory experiments published by these authors were an indication of the type of instruction implemented in the undergraduate laboratory curriculum, since it is likely that these labs were implemented at the authors’ institutions. The level of inquiry of each lab exercise was evaluated through the rubric developed by Fay, Grove, Towns, and Bretz (2007). They found that the majority of the exercises were classified at the lower levels of the inquiry scale: 39% provided the problem, method and solution to the students, and 43% provided both the problem and method to students (DeChenne, et al., 2014). This distribution was similar for exercises published in the American Biology Teacher journal. A longitudinal analysis over the decade investigated across three journals shows maintenance of a status quo.

Summary
The few studies that have been conducted throughout the past two decades all indicate a lack of uptake of inquiry-based instructional practices in the chemistry undergraduate laboratory. Studies on instruction provided
in the laboratory setting are severely lacking in all aspects. There are needs for large-scale studies, studies conducted at various types of institutions and departmental settings for biochemistry (i.e., when biochemistry is part of the chemistry department, the biology department or on its own), and studies that provide insights into similarities and differences in laboratory instructional environments across the undergraduate curriculum.

**ASSESSMENT**

Assessment practices of chemistry and biochemistry instructors have not been extensively studied. A set of studies provides some insight into the knowledge instructors have about formative and summative assessment (Raker, Emenike, & Holme, 2013; Raker & Holme, 2014). Both studies are based on a national survey of about 1,500 chemistry instructors collected during academic year 2009-2010. The authors found that 42% and 44% of those surveyed were aware and had some understanding of the terms formative and summative assessment, respectively (Raker, et al., 2013). This level of knowledge was mostly consistent across the different types of institutions and different sub-disciplines of chemistry (Raker & Holme, 2014).

A national survey of physical chemistry instructors mentioned above, provides insight into the assessment practices in these particular courses (Fox, 2016; Fox & Roehrig, 2015). The authors found that 63% of the respondents (N=331) give a combination of mathematical and subjective (i.e., open-ended) questions even though the majority of instructors had conceptual goals for their courses. This suggests a need for better alignment between learning goals and assessment.

The chemical and biochemistry education research communities have developed different types of assessment tools (cognitive, affective, and content-based) over the last two decades (Bretz, 2013; Holme et al., 2010). A comprehensive list can be found on Dr. Lowery Bretz’s website (Bretz, 2018) and the ASBMB website (ASBMB, 2017c). A national survey of chemistry instructors conducted by Raker and his team in 2017 identified three different assessment profiles among their 784 respondents (Gibbons, Murphy, & Raker, 2018): 45% of respondents relied heavily on written assessments and presentations, 38% on online homework and classroom response systems and 16% on online homework, affective and cognitive surveys and concept tests. Stacey Lowery Bretz, who has developed numerous concept tests, started keeping track of inquiries from instructors for these tests in 2013. As of February 2018, she had received 126 requests from chemistry and biochemistry instructors from around the world (S. L. Bretz, personal communication).

**Assessment of Departments, Colleges, and/or Institutions**

One recent national study explored the program assessment practices of chemistry departments (Emenike, Schroeder, Murphy, & Holme, 2013). Over 1,500 instructors participated in the survey, with an overrepresentation of those from four-year institutions. Instructors were asked whether their department was expected to improve assessment or provide assessment reports. Those who indicated such expectations (72%) were then asked to identify the type of assessment that the department was using, from a list provided. The top three assessments identified were “ACS Standardized Exams (66%), student performance on specific questions (or content) from in-house exams (38%), and student research projects (37%)” (Emenike, et al., 2013, p. 564). Differences existed between the types of institutions: ACS
standardized exams were used more often at four-year and doctorate-granting institutions, while in-house exams were used more often at two-year institutions; student research projects and surveys were more prevalent at four-year and doctorate-granting institutions while assessing student laboratory notebooks was more prevalent at two-year institutions.

Since 2015, ASBMB has required that students in accredited departments take a certification exam, which is aligned with the learning goals of the accreditation program and contains questions across Bloom’s taxonomy (American Society for Biochemistry and Molecular Biology, 2017b). Over 650 students took the certification exam in 2017; over half demonstrated meeting the learning goals described in the accreditation document (Kennelly, 2017).

**Summary**

The assessment practices of instructors and departments are unclear. Moreover, uptake of research-based assessment tools seems limited. The evidence of the uptake of research-based assessment tools is extremely limited. There are needs for large-scale studies, studies conducted at various types of institutions and in varied departmental settings for biochemistry, and across the undergraduate curriculum. Moreover, studies need to address how instructors are implementing these tools and how instructors are leveraging the feedback and information they provide.

**INSTRUCTOR INTEREST, BELIEFS, AND ATTITUDES**

We focus here on reporting the results of studies that investigated instructors’ interest, beliefs, and attitudes related to reform-based teaching. We do not report on studies that focus on characterizing the theoretical relationships between instructors’ belief systems and their instructional practices (Gibbons, Villafañe, Stains, Murphy, & Raker) or studies that explore instructors’ beliefs related to other pedagogical aspects such as curricular goals and content coverage (A. D. Bruck & Towns, 2013; L. B. Bruck, Towns, & Bretz, 2010; Fox & Roehrig, 2015; Kahveci, Gilmer, & Southerland, 2008; Mack & Towns, 2016).

**Instructors**

Stains analyzed data from 126 assistant professors from doctorate-granting institutions from an online survey probing their knowledge, beliefs, and use of RBIS (Pilarz, et al., 2014). Respondents were asked to provide their level of agreement with four statements that provide insights into their thinking about RBIS (Table 3.1). These data indicate that chemistry assistant professors have a positive attitude towards RBIS, although this sample likely overrepresents faculty with above-average interest in teaching. The same questions were asked of chemistry faculty in one department at a doctorate-granting institution (Lund & Stains, 2015), 71% of whom responded in an online survey. The results were strikingly different than those observed with the study of assistant professors. Although they demonstrated the same level of interest in implementing non-lecture strategies (80% agreed or strongly agreed with that statement), 60% to 70% agreed or strongly agreed with the other three items indicating negative views towards RBIS.
Both of these studies also employed the Approaches to Teaching Inventory (ATI) to measure the extent to which chemistry faculty held a teacher-centered versus student-centered attitude. A comparison of results between the two studies indicates similar level for both populations of teacher-centeredness (3.5±0.6 for the assistant professors versus 3.7±0.4 for chemistry faculty within the same department, on a 5-point scale) and student-centeredness (3.5±0.7 for the assistant professors versus 3.3±0.6 for chemistry faculty within the same department, on a 5-point scale).

As part of the evaluation of the Multi-Initiative Dissemination (MID) project—a project funded by NSF to promote the dissemination of RBIS through workshops—Barker and Lewis also collected the ATI from workshop participants (Barker, 2006). Analysis of 203 workshop participants indicated a preference for teacher-centeredness. However, a study has recently demonstrated the lack of reproducibility of the ATI scales, and these results should thus be interpreted with caution (Harshman & Stains, 2017).

Teaching Assistants
Small-scale studies have been conducted on teaching assistants within the context of reforms. These studies provide some insight into TAs’ attitudes and beliefs toward inquiry teaching in the laboratory (Gutwill-Wise, 2001; Kurdziel, Turner, Luft, & Roehrig, 2003; Sandi-Urena & Gatlin, 2013; Wheeler, Maeng, & Whitworth, 2016). Analysis of results from these studies indicates the following common findings:

- TAs’ beliefs toward inquiry teaching depend on their prior experiences as students with this type of practice.
- Professional development programs can help students develop beliefs about teaching aligned with inquiry practices.

Summary
Despite the demonstrated relationships between instructor interest, beliefs, and attitudes towards teaching and learning and instructional practices, we know little about the stance of chemistry and biochemistry instructors on these constructs. Studies on instructor interest, beliefs, and attitudes related to reform-based teaching are severely lacking.

| INSTRUCTOR DEVELOPMENT AND PREPARATION FOR TEACHING |

In the mid-1990s, the National Science Foundation (NSF) funded five projects that aimed to promote active learning in the undergraduate chemistry classrooms (Peace, Lewis, Burke, & Greenbowe, 2002). Each

<table>
<thead>
<tr>
<th>Table 3.1. Chemistry assistant professors’ attitudes toward RBIS. Numbers represent the percentage of assistant professors who agreed or strongly agreed with each statement. (Data from Pilarz, et al., 2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>I am interested in implementing non-lecture strategies in my class</td>
</tr>
<tr>
<td>Teaching with new instructional methods limits content coverage</td>
</tr>
<tr>
<td>Group work is more appropriate in recitation than in lecture</td>
</tr>
<tr>
<td>Teaching with new instructional methods takes more preparation time than lecturing</td>
</tr>
</tbody>
</table>
project reached for this goal by developing new curriculum and instructional practices as well as providing training to instructors who were interested in adopting these strategies. In 2000, NSF funded the Multi-Initiative Dissemination (MID) project to further support dissemination of these projects for a period of four years (Burke, Greenbowe, & Gelder, 2004). Principal investigators of the five projects worked together to design a 1.5-day workshop that introduced chemistry instructors to all five projects. Over the course of the projects, at least 900 people participated in the 34 workshops that were offered (Lewis, 2005). The evaluation of the MID workshops indicated positive impact, with instructors reporting using less teacher-centered strategies and a positive influence on their attitude toward RBIS (Barker, 2006). Some individual projects also obtained additional funding to disseminate their particular strategy (e.g., Process Oriented Guided Inquiry Learning, Peer-Led Team Learning, and Calibrated Peer Review). Workshops to train instructors on these practices are still ongoing today (POGIL, 2017; The Center for Peer-Led Team Learning, 2018). Communications with the POGIL leadership indicated that several thousand undergraduate chemistry instructors have participated in POGIL workshops offered since 2003. More recently, a new workshop project has emerged with funding from the Research Corporation for Science Advancement and the ACS. This project, which started in 2012, focused originally on chemistry assistant professors in their first or second year of appointment at doctorate-granting institutions (Baker, et al., 2014) and has expanded since 2016 to first-, second- and third-year assistant professors at primarily undergraduate institutions due to extended support from the ACS (ACS, 2018a). New faculty participate in a 1.5-day workshop during which they are introduced to RBIS. Since 2012, over 320 faculty have participated. Evaluation of the workshop outcomes indicate a short-term increase in awareness and adoption of RBIS and teaching self-efficacy as well as a shift in beliefs toward student-centered teaching (Stains, Pilarz, & Chakraverty, 2015). The longitudinal aspect of the study, however, indicates difficulties in sustaining these outcomes over time.

In 2015, ASBMB started the Promoting Active Learning and Mentoring (PALM) Network, which was supported by NSF starting in 2016 (PALM Network, 2016). This network is following a mentorship model in which an active learning expert is shadowed by a novice interested in implementing these practices in their classroom. Other experiences offered to fellows include participating in workshops and sharing resources. The aim of this network is to promote active learning. To date, 13 fellows are part of the network. Other types of professional development opportunities have been provided to chemists and biochemists. These workshops typically focus on the development of instructional materials. For example, both the ACS, through its Exam Institute, and ASBMB have designed and implemented workshops with the aim of engaging instructors in the writing of test items for their respective national exams. Several workshops for chemists focus on the development of new content materials that incorporate research-based assessment and learning strategies (Interactive Online Network of Inorganic Chemists, 2018) or training instructors in new content areas (cCWCS, 2018).

**Professional Development for Graduate Students**

The nature and extent of professional development provided to chemistry
graduate students is largely unknown. Numerous studies have described innovative professional development programs and their impact on TAs (Dragisich et al., 2016; Dragisich, Keller, & Zhao, 2016; Mutambuki & Schwartz, 2018; Pentecost, Langdon, Asirvatham, Robus, & Parson, 2012; Wheeler, Clark, & Grisham, 2017; Wheeler, et al., 2016) but the extent to which these programs have been propagated is unknown.

Resources
Members of the chemical education research community put together in the mid-late 2000 two books intended to provide guidance to chemistry instructors about research-based teaching practices and the latest research developments on student learning in chemistry (Pienta, Cooper, & Greenbowe, 2005; Pienta, Cooper, & Greenbowe, 2009). These books are now out of print.

Summary
Professional development programs targeting instructors seem to be more nascent in biochemistry and thus their impact has not been fully established. In chemistry, there are some very active professional development programs linked to specific RBIS, and a growing focus on broader training of the professoriate, especially by the ACS. The pedagogical training that teaching assistants receive is unclear in both fields.

Most of the evidence on the level of engagement of instructors in professional development workshops comes from the facilitators of these workshops. The impact of national professional development program on participants has not been systematically investigated. We do not have evidence to provide a national picture of the type (i.e., local or national, focus on one RBIS or multiple, length of program) and the number of professional development activities that chemistry and biochemistry instructors engage in. We also do not know who engage when in such opportunities.

DIVERSITY AND INCLUSION

The Higher Education Research Institute (HERI) national surveys, described earlier, provide the only national outlook on the inclusion of inclusive practices that we could find. Eagan analyzed the HERI surveys from 2014, when an item specific to inclusive practices was first included. This item asked faculty to indicate the extent of their use of “techniques that create an inclusive classroom for diverse students” (Eagan, 2016). He found that only 16.0% of the physical sciences faculty and 22% of the biological sciences faculty indicated using these strategies in all of their courses, while 38.7% of physical science faculty and 28% of the biological sciences faculty reported not using these strategies in any of their courses.

The American Chemical Society (ACS) hosted three workshops in the 2000s that focused on identifying issues and solutions to increasing the number of African American, Native American, and Hispanic students in the chemical sciences. Each of these workshops led to a report that laid out recommendations (ACS Committee on Professional Training, 2004; ACS Committee on Professional Training & ACS Committee on Minority Affairs, 2008a, 2008b).

Summary
There seems to be little momentum on this issue. This area is ripe for investigations, as little work has been done.

CURRICULUM RESOURCES, DEVELOPMENT, PLANNING AND USAGE

In the mid 1990s, NSF funded five projects that aimed to promote active learning in the undergraduate chemistry classrooms
Each project reached for this goal by developing new curriculum and instructional practices. We report here on their current uptake of the products of these projects based on information that we could collect.

**Process Oriented Guided Inquiry Learning (POGIL).** In a POGIL classroom, students work cooperatively on worksheets that have been designed around the learning cycle. The instructor serves as a facilitator. More information about POGIL and resources for implementation can be found on the POGIL website (POGIL, 2017). POGIL leaders provided insight into the propagation of this strategy. Based on their latest survey, at least 303 chemistry instructors are using POGIL (B. Miller, personal communication). The characteristics of these users and the course they teach are provided in Table 3.2. Leaders indicated that these numbers underestimate POGIL uptake, as many users do not fill out their surveys.

**Peer-Led Team Learning (PLTL):** This strategy consists of engaging small groups of students in collaboratively solving problems for 1.5 to two hours per week. Each small group is guided by a peer leader, a student who recently passed the course with a high grade. More details about the strategy and resources for its implementation can be found in the PLTL website (The Center for Peer-Led Team Learning, 2018). In a recent report from PLTL leaders, it was indicated that over 150 institutions were using PLTL in the US, with distribution across the different types of institutions (Wilson & Varma-Nelson, 2016). Moreover, it was estimated in 2008 that over 20,000 students, 150 instructors and 1,500 peer leaders are engaged in PLTL each year (Gafney & Varma-Nelson, 2008). A national sample of chemistry instructors was surveyed in 2017 about their instructional practices, which included PLTL. In this study, the authors found that one in four instructors in their sample (N=829; 14.4% response rate) reported implementing PLTL in their course multiple times throughout a semester (Srinivasan, Gibbons, Reed, Murphy, & Raker, 2018). The authors found no difference in implementation rates based on institution types and faculty ranks. Finally, they found that PLTL was least implemented in small, upper-level courses.

**Calibrated Peer Review (CPR):** This instructional management tool was developed within chemistry but has since been used in STEM and non-STEM disciplines.

### Table 3.2. Characteristics of POGIL users. (Data from B. Miller, personal communication)

<table>
<thead>
<tr>
<th>Type of institution</th>
<th>Chemistry discipline</th>
<th>Frequency of use</th>
<th>Use of official POGIL materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-year Colleges</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18%</td>
<td>Non-Majors</td>
<td>Daily</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Preparatory Chemistry</td>
<td>2-3 times per week</td>
<td>59%</td>
</tr>
<tr>
<td>4-year Colleges</td>
<td>General, Organic,</td>
<td>Once a week</td>
<td>No</td>
</tr>
<tr>
<td>45%</td>
<td>Biochemistry</td>
<td>19%</td>
<td></td>
</tr>
<tr>
<td>University</td>
<td>General Chemistry</td>
<td>2-3 times per month</td>
<td>12%</td>
</tr>
<tr>
<td>37%</td>
<td>Organic</td>
<td>Once a month</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Physical</td>
<td>Less than once a month</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Inorganic</td>
<td>Once per semester</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analytical</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Biochemistry</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
disciplinary. The tool allows instructors to promote conceptual understanding by engaging students in writing assignments and peer-review of these assignments. According to the CPR website (2017), by 2011, the tool had been adopted “in over a thousand institutions, in undergraduate and graduate programs, in professional medical and business schools, and even in secondary schools. Faculty in small private schools through large state universities have integrated CPR assignments in over 5000 courses.”

ChemConnections: The aim of this project was to “to develop new curricula, materials and methods to enhance the appreciation and learning of science, especially chemistry, for every undergraduate student such that all college graduates will command the knowledge and skills necessary to permit continued learning, to lead productive lives, and to make informed decisions” (ChemConnections, 2007; Anthony, Mernitz, Spencer, Gutwill, Kegley & Molinaro, 1998). The result is a series of modules that teach chemistry principles within relevant real-world contexts. The modules are now published by W.W. Norton and Company, Inc. and have been updated since and incorporated into a student workbook. Over 50% of faculty who attended workshops promoting these materials adopted them in their courses (Lewis & Lewis, 2006).

New Curricula for General Chemistry
Chemistry, Life the Universe and Everything (CLUE). This new general chemistry curriculum developed and initially tested in 2010 is grounded in research on teaching and learning. It is centered on four crosscutting concepts: structure and properties, bonding and interactions, energy, and change and stability. More details about the curriculum and access to resources and publications can be obtained from the CLUE website (Cooper & Klymkowsky 2016). To date, CLUE has been implemented at four different institutions and has involved over 13,000 students and multiple instructors (S. Underwood, personal communication).

Chemical Thinking: This new general chemistry curriculum, also based on research on teaching and learning, focuses on helping students recognize chemistry as a powerful way of thinking that can help them understand real world problems facing our time. For more information about this curriculum and to access resources, please check their website (Talanquer & Pollard, 2018) and publications (Kober, 2015; Sevian & Talanquer, 2014; Talanquer & Pollard, 2010). To date this curriculum has been fully adopted at the institution where it was designed.

Chemistry in Context: In 1989, the ACS embarked on a project aimed at developing a textbook about chemistry for non-science majors. The result was the textbook called Chemistry in Context and teaches chemistry principles through real-world issues (Middlecamp, 2008). It is currently in its 9th edition. According to a representative from the ACS, Terri Taylor, this is one of the most popular chemistry textbooks for non-STEM majors on the market. The ACS disseminates it through workshops hosted at conferences but also through on campus visits. These on-campus workshops are typically conducted by members of the author team.

General chemistry performance expectations: Over the last two years, the ACS has hosted workshops on general chemistry performance expectations (Wang, 2016). The goal of this new initiative is “to advance the development of general chemistry performance expectations so as to better align course design, instruction, assessment, and student transfer across and within higher
education institutions.” The workshop targets two-year and four-year institutions.

**Biochemistry Curriculum**

As part of the building up of the accreditation program, ASBMB conducted workshops with biochemistry instructors to help establish the core ideas, underlying theories and foundational skills that undergraduate biochemistry majors must have (Tansey et al., 2013; White, Benore, Sumter, Caldwell, & Bell, 2013; Wright, Provost, Roecklein-Canfield, & Bell, 2013). Materials align with these learning goals are starting to be populated in CourseSource, a peer-reviewed repository of research-based teaching resources (CourseSource, 2018).

**Inquiry Curriculum in the Laboratory**

A study conducted by Buck, Bretz, and Towns (2008) in the late 2000s analyzed the level of inquiry provided by 13 published general and organic chemistry laboratory manuals. These manuals were specifically selected because they either included the word “inquiry” in the title or made reference to the literature on inquiry. The authors leveraged prior literature on inquiry to design a rubric that characterizes the extent to which students are provided with the problem, background and theory, methods, how data should be analyzed, how data and results should be presented, and what should have been observed and concluded. They found that 83% of the experiments provided students with structured inquiry experiences, in which all information is provided to students except for how data and results should be presented and what should have been observed and concluded. Only 9% of the experiments were classified as guided inquiry, 2% as open inquiry, and none as authentic inquiry, levels that provide less scaffolding and are more demanding for students.

A major program reform for the laboratory was led by the Center for Authentic Science Practice in Education (CASPiE) for a couple of years in the early 2000s. The goal of this program was to provide authentic research experiences to students enrolled in the first and second year of chemistry courses. The program consists of engaging research faculty in the development of laboratory modules that contribute to their research program. These modules are then implemented in the gateway chemistry courses and the research is carried out by students in these courses. According to the following website (Association of American Universities, 2018), CASPiE has engaged 6,000 students across 17 higher education institutions of different types.

CASPiE’s model is similar to what is referred to as CURE, or Course-based Undergraduate Research Experience. The implementation of CUREs seems to be on the rise in chemistry, based on the number of submissions of laboratory articles to the *Journal of Chemical Education* that focus on a CURE (R. Cole, personal communication), and already active in biochemistry, based on the number of articles published in *Biochemistry and Molecular Biology Education*. Within biochemistry, communities promoting the propagation of CURE experiences have emerged over the last couple of years (BASIL, 2015; Bell et al., 2017; Craig, 2017).

**Other Instructional Tools**

PhET simulations are interactive computer simulations to help students understand STEM concepts. They are freely available to students and instructors. To date 60 simulations focused on chemistry concepts have been developed. PhET leaders indicated that 4,569 undergraduate chemistry instructors were registered on the site; however they indicated that the majority of chemistry users do not register (E. Moore, personal communication).
A group of inorganic chemists came together in 2006 to start developing a community of inorganic chemists interested in developing and using “learning objects” or pieces of curriculum. One outcome of this process is a website hosting these learning objects along with other resources (IONIC, 2018). As of March 2018, the site hosted 860 learning objects that could be accessed by over 1,200 registered instructors, both national (every state in the US as at least one user) and international (about 40 countries) (Stewart, 2018). Leaders shared with us the result of a survey they conducted in 2016. They contacted 800 registered users of the site and 100 responded. Table 3.3 below provides demographic data about these respondents. Seventy-six percent of the survey respondents say they use VIPEr learning objects in their teaching.

**Content Coverage Studies**

Over the last few years, there has been a great interest within the chemical education research community (primarily led by members of the ACS Exam Institute) to characterize the content taught at various levels of the undergraduate chemistry curriculum (Fox & Roehrig, 2015; Holme, Luxford, & Murphy, 2015; Holme & Murphy, 2012; Holme, Reed, Raker, & Murphy, 2017; Luxford & Holme, 2015; Marek, Raker, Holme, & Murphy, 2017; Raker, Holme, & Murphy, 2013; Raker & Holme, 2013; Raker et al., 2015a, 2015b; Reisner, et al., 2015). This set of papers, although not reporting on the use of RBIS, provides useful insight into what is being taught in the undergraduate chemistry curriculum.

**Summary**

A significant body of active learning curricular materials has been developed in chemistry and is starting to be developed in biochemistry. Some of these resources are more used than others. Data collected to date, mostly by the designers of the RBIS or publishers of curriculum materials, focus on number of users. There are few studies exploring how the materials are actually being implemented in classrooms.

**Table 3.3. Characteristics of users of VIPEr learning objects.**

<table>
<thead>
<tr>
<th>Type of institutions</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Institutional control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>Public</td>
<td>62</td>
<td>51</td>
</tr>
<tr>
<td>Highest chemistry degree awarded</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s Degree</td>
<td>92</td>
<td>91</td>
</tr>
<tr>
<td>Master’s Degree</td>
<td>24</td>
<td>24</td>
</tr>
<tr>
<td>Doctoral Degree</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Associates Degree</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Institutional size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than 500 students</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>500 to 1,500 students</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>1,500 to 5,000 students</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>5,000 to 15,000 students</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td>15,000 to 30,000 students</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Greater than 30,000 students</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Research versus teaching mission</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research Intensive</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Liberal Arts</td>
<td>54</td>
<td>53</td>
</tr>
<tr>
<td>Comprehensive</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>Predominantly Undergraduate</td>
<td>55</td>
<td>55</td>
</tr>
<tr>
<td>Community College</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Other: [Predominantly minority]</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

(Data from survey conducted by VIPEr leaders, shared with author. See https://www.ionicviper.org/)
FACTORS AND PROCESSES

Several studies have investigated factors influencing the propagation of RBIS in STEM (Kezar & Gehrke, 2015; Khatri, Henderson, Cole, & Froyd, 2013; Khatri et al., 2016; Khatri et al., 2017; Stanford et al., 2015; Stanford et al., 2017). Chemistry RBIS were included in these investigations. We summarize below the findings from these studies along with findings that emerged from other readings and personal communications.

FACTORS ENABLING PROPAGATION OF CHEMISTRY REFORMS

One of the main events that led to the growth in development, dissemination, and propagation of RBIS in chemistry is the launch of the Systemic Changes in the Undergraduate Chemistry Curriculum (Chemistry Initiative) in 1994. This program permitted the funding of five projects, several of which are still extremely active today and have seen widespread propagation. Elements thought to contribute to this success included NSF’s requirements for each project to include other institutions than those leading the project, and for the five projects to conduct reciprocal site visits that involved multiple leaders from each project. These two requirements enabled the development of a large community of chemistry instructors, across many types of institutions, who were interested in similar goals. Moreover, NSF also funded in the late 1990s nine ‘Adaption and Adoption’ awards to encourage uptake of these approaches. Susan Hixon, the NSF program director from Chemistry DUE from 1992 to 2012, also indicated that the success of these efforts came from a small group of committed instructors and leaders who recognized the need to assemble a community that would support the dissemination, support of adopters, hosting of workshops, and gathering of funding (NRC, 2014).

Overlapping with these events was the growth of Project Kaleidoscope (PKAL). This project which was initiated in 1989, aims at developing and supporting a network of faculty and administrators towards improving the education of STEM students. Several of the people we talked to commented on the impact that PKAL had in shaping the national conversation about STEM undergraduate teaching in the 1990s and helping to legitimize student-centered teaching.

In biochemistry, the *Vision and Change* document guided the development of the learning goals for the accreditation program.

NSF funded in 2011 a grant project called *Increase the Impact* aimed at characterizing effective practices in the propagation of STEM RBIS, which included POGIL and PLTL (Increase the Impact, 2015). The research team analyzed propagation strategies of projects supported by the NSF-TUES program, and interviewed principal investigators and NSF program directors. Their efforts led to the development of a model and guide to assist future grant awardees in planning for successful propagation of their research products (Henderson et al., 2015; Khatri et al., 2013; Khatri et al., 2016; Stanford et al., 2015). The propagation model that emerged from this research is presented in the figure below (from Khatri et al., 2016).

Take-aways from this work (Khatri et al., 2016; Stanford, et al., 2017) include:

- The importance of understanding potential adopters and their instructional systems. Developers of successfully propagated RBIS need to identify during the development phase potential adopters, their needs, and the personal and contextual constraints they may experience when implementing the RBIS.
The value of engaging potential adopters. Developers of successfully propagated RBIS engage potential adopters during the development and testing phase of the RBIS.

The need to identify design principles. Developers of successfully propagated RBIS identified through various rounds of testing and in different environments, the design principles of their RBIS, i.e., what are the key features of the strategy that contribute to the positive outcomes? Typically, RBIS that are more adaptive and less prescriptive were more successful.

The importance of using interactive dissemination strategies. Successful developers implemented both passive dissemination strategies (e.g., publishing papers, presenting at conferences, posting on websites) and interactive strategies that include interactions between developers and potential adopters (e.g., workshops) from the beginning of the project. Moreover, developers designed their dissemination strategies based on their beliefs about how change takes place and adapted their strategies based on their analyses of reasons behind low uptake of their RBIS.

The challenges of providing follow-up and support. Developers of successfully propagated RBIS developed mechanisms to support adopters, such as developing how-to-guides, conducting site visits, and hosting workshops.

Successful projects received continuous funding over an extended period of time.

Kezar and Gehrke conducted an extensive study on four STEM communities of practices (CoPs) focused on pedagogical reform which were extremely successful in propagating their strategies (Gehrke & Kezar, 2016, 2017; Kezar & Gehrke, 2015, 2017). POGIL was one of the CoPs. They were interested in identifying common characteristics across these CoPs that led to STEM reform in the CoP participants’ department and institution. Their work shows that POGIL and the other CoPs have been successful because of a set of defining characteristics that include inclusion, adaptation, and leadership development, similar to the effective propagation approaches emphasized by Khatri and coauthors.

The CoPs help their participants develop leadership skills and become change agents.

The CoPs have a philosophy of teaching and reform that guides the work and that participants abide to.

Fig. 3.1: Propagation model

The CoPs identify areas of strength—called leverage points—and build on them; for example, one of POGIL’s strengths was its disciplinary focus.

The CoPs build leadership at different levels of their network and ensure continuity of leadership through identification and nurturing of potential future leaders.

The CoPs continuously obtain feedback from various stakeholders to improve project.

The CoPs conduct research and assessment of their practices and leverage results from these studies to recruit new participants and secure funding.

The CoPs have a community-level strategic plan that provide clear directions for the community and support its growth.

The ASBMB has created a CoP around curricular reform that seeks to meet many of these criteria; this may partly explain the rapid uptake of their accreditation program.

FACTORS INHIBITING OR SLOWING PROPAGATION

Several chemistry-specific factors seem to have contributed to the slow uptake of RBIS. First, chemistry education researchers (CER) had to focus in the 1980s and 90s on establishing the legitimacy of their field among the chemistry professoriate and research community (Bodner, 2011; Herron & Nurrenbern, 1999; Webb, 2007). As a consequence, research questions were more focused on disciplinary content and student outcomes. Interest in investigating the uptake of RBIS has only grown over the last few years. To date, CER scholars still struggle to have their work published in ACS journals focused on traditional bench chemistry (NRC, 2014).

Second, communities are not interacting as much as they should. There is limited interaction between chemists and CER specialists at the ACS national meetings (NRC, 2014). Members of the biochemistry education researcher community have split alignment between either biology education research or chemical education research, which limits their progress.

Third, although the accreditation documents from the ACS evolved over time toward supporting RBIS, resources to help departments and instructors implement these recommendations are limited (Towns, 2009). ASBMB has been much more proactive in this respect and provides resources within their accreditation guide (ASBMB, 2017a).

Fourth, some studies suggest that many (not all) chemistry instructors have a mindset regarding teaching that potentially limits progress (Emenike, et al., 2013; Lund & Stains, 2015).

Fifth, the assessment tools used to measure the impact of reform need to be better aligned with expected outcomes of such reforms, which means moving beyond the content and measuring skills and attitudes that all chemists and biochemists value (Anderson, 2007; Holme, et al., 2010; Holme, Cooper, & Varma-Nelson, 2013).

Sixth, except for PLTL, few studies have identified critical components of RBIS that translate into positive outcomes in various contexts (Holme, et al., 2013; Stains & Vickrey, 2017). This significantly diminish the propagation process, as potential adopters may not see expected outcomes when first implementing RBIS (Chase, Pakhira, & Stains, 2013) and as a result may be more likely to drop it.
Finally, we are still missing studies that provide a tested framework for instructional innovation within the realm of chemistry and biochemistry and across institution types (Holme, et al., 2013). The lack of framework makes it difficult for developers of RBIS, professional societies, and departments or faculty interested in transforming teaching locally to plan and implement effective change strategies.

The Increase the Impact project identified several other factors that limited propagation of the RBIS they studied (Khatri, et al., 2013; Khatri, et al., 2016):

▲ RBIS developers feel ill-equipped to disseminate their research; they often start planning dissemination after the RBIS is developed and focus on passive dissemination.

▲ Evidence of efficacy for many RBIS is still missing, which may diminish the quality of the argument used to convince potential adopters. For example, NSF program directors interviewed suggested that evidence be collected across a range of institutions, not just at the developers’ institution. Holme, Cooper and Varma-Nelson (2013) suggested that more longitudinal studies were need in order to provide evidence and an understanding of the long-term impact of the implementation of a reform.

▲ Disciplinary societies can play a bigger role in advocating for integration of RBIS in curricula.

Summary
Much of the data that we were able to gather on the uptake of specific RBIS came from individuals themselves rather than publications. It is thus critical to ensure evaluators of projects measure uptake of practices and publish these results or make their evaluation report public. Surveys and classroom observations have been useful in providing large-scale adoption patterns, but each method has its limitations

FUTURE DIRECTIONS
This analysis reveals several shortfalls in the evidence available to characterize the uptake of RBIS.

▲ Large-scale national studies across institutional types are necessary. Most studies seem to be currently conducted at doctorate-granting institutions. We do not know how and to what extent RBIS are being propagated at other institution types.

▲ Studies investigating the uptake of RBIS need to include fidelity of implementation measures (Stains & Vickrey, 2017). For almost all RBIS reported here, we do not know how instructors are implementing them. There is a large difference between reporting the number of users and knowing how these users are actually implementing the strategies.

▲ We need to investigate and characterize factors that enhance propagation within the context of a discipline and make comparisons across disciplines. Currently, instructional change studies in DBER and higher education include all fields together. Comparison of trends in uptake between chemistry and biochemistry and studies comparing factors influencing instructors in different types of STEM disciplines (Lund & Stains, 2015) point to a potential disciplinary effect. If this effect exists, a universal propagation framework may not be appropriate.
We are missing information about the contexts in which these RBIS and active learning could be implemented. Most pedagogical and content RBIS focus on the lecture portion of the course. Consequently, this is where most measurements have been made. We do not know the extent to which students experience RBIS and active learning in the laboratory, recitation, or field.

REFERENCES


Bell, J. K., Eckdahl, T. T., Hecht, D. A., Killion, P. J., Latzer, J., Mans, T. L., … Ellis Bell, J. (2017). CUREs in biochemistry—where we are and where we should go. Biochemistry and Molecular Biology Education, 45(1), 7-12.


cCWCS (2018). Chemistry Collaborations, Workshops and Communities of Scholars Retrieved 03/27, 2018, from http://www.ccwcs.org/content/about-ccwcs


graduate teaching assistants’ teaching. Journal of Chemical Education, 83(2), 305.


Adoption of Research-based Instructional Strategies in Engineering and Computer Science

Cynthia J. Finelli, University of Michigan and Maura Borrego, University of Texas at Austin.

INTRODUCTION

Extensive research has shown that, compared to traditional, lecture-based teaching practices, research-based instructional strategies (RBIS) can improve student learning, engagement, and success in STEM (e.g., Crouch & Mazur, 2001; DesLauriers, Schepere, & Wieman, 2011; Freeman et al., 2014; Prince, 2004; Springer, Stanné & Donovan, 1999). However, in spite of this evidence, translation of research to actual classroom practice has been slow (Jamieson & Lohmann, 2012; NRC, 2012; PCAST, 2012; Stains et al., 2018), and many engineering and computer science (CS) classrooms feature mostly traditional teaching practices. They use lecture-based pedagogy, focus on theory rather than practical applications, employ a competitive grading system, emphasize algorithmic problem-solving, and are designed to challenge students who learn best in that environment while providing limited (if any) support for capable students who learn differently (Bransford, Darling-Hammond, & LePage, 2005; Brawner, Felder, Allen, & Brent, 2002; Hora, Ferrare, & Oleson, 2012).

RESEARCH-BASED REFORMS

Like many other STEM fields, scholars in engineering and CS have attempted to measure the extent of RBIS use in undergraduate classrooms across the United States. Much of the early work was informed by a diffusion of innovations perspective (Rogers, 2010), and, thus, focused on stages of adoption, barriers, and sources that can influence instructors to use RBIS.

Borrego and her colleagues surveyed instructors to ascertain the extent to which STEM instructors had adopted RBIS (Borrego, Cutler, Prince, Henderson, & Froyd, 2013; Borrego, Froyd, & Hall, 2010; Froyd, Borrego, Cutler, Henderson, & Prince, 2013; Henderson, Dancy, & Niewiadomska-Bugaj, 2012; Prince, Borrego, Cutler, Henderson, & Froyd, 2013). Those data demonstrated that instructors drop out of the adoption/innovation cycle at various points, and...
more instructors initially try RBIS and then abandon their use than never use them.

Figure 4.1 presents data from electrical, computer, and chemical engineering instructors across the U.S. More than one-third of instructors, the largest group, discontinued their use of RBIS after trying three or more strategies. This is an important finding because so much effort has been on building basic awareness of RBIS and convincing STEM instructors to try them. Yet it appears that sustaining RBIS use is a bigger issue than building awareness or encouraging an initial trial.

In a different analysis of the same survey data, Froyd and colleagues (Froyd et al., 2013) presented results from 115 electrical and computer engineering instructors who had recently taught a “sophomore-level engineering science course (circuits, electronics, or introductory digital logic and/or digital design).” In engineering and CS, these sophomore-level courses are often the initial entry point in the major, and since they are content-heavy and often conceptually difficult, they are a prime target for RBIS use. The researchers’ analysis focused on proportions of instructors who (1) were aware of RBIS but never implemented them; (2) tried RBIS previously but abandoned them; or (3) were currently using RBIS in these sophomore-level courses. Current use at the time of the survey was particularly high for “active learning” (62% of respondents) and “collaborative learning” (51%), while abandonment was particularly high for Just-in-Time Teaching (66% of instructors who ever tried it), thinking aloud/paired problem solving (55%), think pair share (54%), and problem-based learning (53%).

The researchers reported very similar results for 92 chemical engineering instructors surveyed during the same period (Prince, Borrego, Henderson, Cutler, & Froyd, 2013). Nearly two thirds of instructors reported current use of active learning (65%) and collaborative learning (60%), and approximately one third each currently used PBL (35%) and inquiry-based learning (31%). Again, the analysis focused on information channels, including how often and with whom instructors discussed their teaching.

A separate study provides some insight into CS adoption rates of RBIS. In 2010, Borrego, Froyd and Hall (2010) surveyed engineering and CS department chairs about awareness and adoption of RBIS, such as learning communities and interdisciplinary capstone design courses, that would require support
from the department chair. Seventy-one percent of all chairs surveyed reported adoption in their departments of “student-active pedagogies” (basically RBIS). For CS and software engineering chairs (in colleges of engineering), the adoption rate was 61%. Design projects in first-year courses were reported to be adopted by 65% of all departments and 25% of computing departments. Open-ended responses to this survey emphasized the value of disciplinary societies and disciplinary networks (i.e., targeting each branch of engineering as a separate network) in diffusing information and influence about RBIS.

A recent study involving classroom observations from over 2,000 classes taught by more than 500 STEM instructors across 25 institutions presents similar findings (Stains, et al., 2018). In both engineering (159 classroom observations) and CS (61 classroom observations), over 50% of the observed classes featured primarily didactic lecturing, while only 15% used student-centered teaching (i.e., RBIS).

**OUR DATA REGARDING RBIS UPTAKE**

Our own research to investigate instructors’ concerns about student resistance offers a snapshot of the types of RBIS in current use in engineering. We surveyed both instructors and their students about active learning and observable student resistance behaviors. Because it is difficult to reliably measure whether students are cognitively engaged, the items on our instructional items survey scale focus on what students were doing during class time (e.g., solving problems, working in groups). One of our publications focuses on the difficulty of creating a factored scale for measuring RBIS instruction (Nguyen, et al., 2017a), and another describes our findings about RBIS use among the instructors who volunteered for our study (Nguyen et al., 2017b). Although these instructors employed substantial RBIS in their teaching, a majority of class time was spent on more traditional approaches in which students “listen to the instructor lecture during class” and “watch the instructor demonstrate how to solve problems during class.” Average use of these two teaching activities was once a week or more, while average use of all other RBIS was “sometimes” (5-10 times per semester).

The very definition of RBIS may differ across engineering and CS disciplines, as compared with other STEM disciplines. Design experiences, team projects, and internships, for instance, are no longer considered RBIS in engineering, for instance. Design experiences in the first and final years of undergraduate engineering curricula were promoted heavily through the 1990s, and they are now a common feature of most engineering degree programs (Froyd, 2005). These design experiences are often completed as team projects, which are also widespread in engineering education (Borrego, Karlin, McNair, & Beddoes, 2013). Similarly, internships and co-ops, where students work part- or full-time for a private company or government organization during the semester or summer, are so common in engineering that there is little published on the practice.
Other STEM disciplines, however, are working to develop these opportunities for their students and may be framing their use as RBIS.

On the other hand, engineering educators have not done much work in developing course-based undergraduate research experiences. While the NSF Research Experiences for Undergraduates (REU) program is as common a funding mechanism in engineering as it is in other STEM fields, engineering educators have innovated primarily in developing REU sites to serve particular student populations (e.g., students with attention-deficit disorder or students with physical disabilities). Engineering degree requirements typically include several lab courses, but to date there is almost no innovation or research in laboratory pedagogy beyond development of activities and modules to teach very specific concepts.

**THE ACCREDITATION BOARD FOR ENGINEERING AND TECHNOLOGY**

One controversial driver of undergraduate educational change in engineering, and to a lesser extent in CS, is accreditation. Among STEM disciplines, engineering uniquely incorporates accreditation through the Accreditation Board for Engineering and Technology, Inc. (ABET), and similar organizations in other regions around the world, as an integral part of the landscape for educational change. ABET comprises four commissions: engineering, computing, engineering technology, and applied science. Adoption of outcomes-based assessment by ABET around 2000 was a major shift for engineering education (Froyd, Wankat, & Smith, 2012; Volkwein, Lattuca, Harper, & Domingo, 2007), which generally preceded similar regional accreditation requirements. However, in recent years, engineering accreditation has become routinized, and the path of least resistance is to keep doing everything the same way. Borrego and Henderson (2014) conclude, “On the one hand, it has positively impacted evidence-based assessment and teaching in engineering; on the other hand, ABET has been criticized for stifling instructional innovation and improvement in engineering education” (p. 237).

One criticism of ABET’s current requirements is that none of the measures encourage innovation in professional development. For example, some have suggested that the criterion about instructor qualifications should require some evidence of faculty development or teaching improvement efforts, rather than simply relying on educational background and/or industry experience. The new ABET criteria, with important changes to requirements related to student outcomes and curriculum (www.abet.org/accreditation/) will hopefully encourage more innovation and have a positive impact on adoption of RBIS in engineering and CS.

**WHY DO INSTRUCTORS ADOPT/NOT ADOPT RBIS?**

Various reasons have been hypothesized for the slow diffusion of innovation and the lack of sustained use of RBIS. For instance, among science instructors, Handelsman and colleagues (2004) noted that many instructors are unaware of the data behind RBIS, distrust the data, or feel intimidated by the challenge of learning new methods. We believe the same applies to many engineering and CS instructors. Many researchers have studied instructors’ motivation to adopt RBIS, and they have identified factors that influence adoption and continued use of RBIS (Barker & Gruning, 2014; Barker, Hovey, & Gruning, 2015; Borrego, Froyd, & Hall, 2010; Dancy & Henderson, 2010; Finelli, Daly, & Richardson, 2014; Fossati & Guzdial, 2011; Friedrich,
Sellers, & Burstyn, 2007; Froyd, et al., 2013; Hora, Ferrare, & Oleson, 2012; Jamieson & Lohmann, 2012; Prince et al., 2013; Seymour, DeWelde, & Fry, 2011; Sunal, et al., 2001). Among the factors that inhibit adoption, lack of time has frequently been raised as one of the most common. Other barriers include fear of student resistance, lack of familiarity with RBIS, lack of skills and knowledge, lack of resources and support for instructors, restrictive course syllabi and content structure, physical classroom limitations, institutional policies (especially as related to tenure and promotion), institution type and research emphasis, teaching evaluations, heavy workload, and misaligned reward systems. Interestingly, according to one study (Henderson, Dancy, & Niewiadomska-Bugaj, 2012), “faculty age, institutional type, and percentage of job related to teaching were not found to be barriers to knowledge or use at any stage” (p. 020104-1). We similarly found that instructor and course characteristics alone do not explain student reactions to RBIS (Nguyen, et al., 2017c; Finelli et al., 2018).

WHAT CAN WE DO TO PROMOTE GREATER ADOPTION OF RBIS?

Professional development for engineering and CS instructors has been shown to be an effective mechanism for promoting evidence-based change, both at the institutional level and more broadly. At the University of Michigan (UM), researchers developed and implemented the Teaching Circle to overcome barriers to adoption and motivate instructors’ change (DeMonbrun, Canas, & Finelli, 2018; Finelli, Daly, & Richardson, 2014). That faculty learning community program was grounded in theory about instructors’ motivation, RBIS, and professional development programs, and it applied a local lens to the national research by collecting concrete data about teaching practices of engineering instructors and student perspectives about supportive teaching at UM. The data pinpointed RBIS that would be most suitable to emphasize with instructors in the UM context, and that would have the greatest potential to impact their students’ success.

We note that this approach of local, STEM-specific faculty learning community support, sustained over at least one academic term, is currently valued as a best practice for supporting use of RBIS. Since its inception at UM in 2012, more than 50 engineering instructors have participated in the Teaching Circle program, and it has had a positive impact on the culture of teaching and learning in UM’s college of engineering. Participants’ enthusiasm, clarity, and interaction increased significantly (while these indicators decreased for instructors in a control group). In addition, the researchers found through classroom observations that many participants changed their teaching to increase student engagement and active learning, and longitudinal studies of student evaluations of teaching indicated that participants in the program had significantly more positive teaching evaluations relative to a control group (Anderson & Finelli, 2014; DeMonbrun, Kerst, Pfershy, & Finelli, 2018).

National professional development programs for engineering and CS instructors have also been successful at influencing instructors’ behavior. One report on the impact of the faculty development efforts of the Southeastern University and College Coalition for Engineering Education (SUCCEED) suggests that well over half of 509 coalition instructors who completed an online survey five years after the coalition began persisted in using RBIS (Brawner, et al., 2002). Most of the respondents attributed their RBIS use to participating in teaching workshops and seminars offered through SUCCEED. Another
broad effort, the National Effective Teaching Institute (NETI), has been offered annually since 1991, and more than 1,000 engineering instructors have attended. A systematic analysis of the impact of NETI (Felder & Brent, 2010), grounded in the theories of adult motivation, demonstrates the wide success of that program. Participants have consistently rated the program as excellent or good, they report adopting (and continuing to use) RBIS highlighted in the program, and they report receiving increased student ratings of teaching.

WHERE IS THE FUTURE?

Perhaps one of the most promising directions for more widespread adoption for RBIS lies in professional development for future engineering and CS instructors. Programs that prepare graduate students for their future instructor roles often make the use of RBIS the status quo. The Center for the Integration of Research, Teaching, and Learning (CIRTL, www.cirtl.net/) and the University of Michigan (UM) Postdoctoral Short Course on College Teaching in Science and Engineering (www.crlt.umich.edu/programs/psc) are two examples. They both offer programs to prepare graduate students to use RBIS as future instructors, and they both have been shown to have a lasting effect. The CIRTL programs, for instance, have been shown to promote reflective teaching practice and develop a learning community of scholars interested in improving STEM education (Austin, et al., 2009; Micomanaco, 2011). And the UM postdoc program has been offered 12 times (in both face-to-face and purely online formats) to more than 400 postdoctoral scholars in engineering and science. The participants found the program to be extremely valuable (average rating of the overall value of the program = 4.6/5.0).
Issues related to supporting a diverse population of students continue to persist in engineering and CS. Women have historically been underrepresented in both engineering and CS, and national statistics show that the share of bachelor’s degrees awarded to women in engineering has not changed in the past 15 years, remaining steady at 20%. In CS, women’s share has suffered a serious decline, dropping from 28% in 2000 to just 18% in 2015 (NSB, 2018).

Hispanics, Asians/Pacific Islanders, Black/African-Americans, and American Indians/Alaska Natives also continue to be significantly underrepresented in engineering and CS. There has been a slight increase in the share of bachelor’s degrees awarded to Hispanics (from 7.4% in 2000 to 10.9% in 2015 in engineering, and from 6.2% to 10.4% in CS), and in the share awarded to Asians/Pacific Islanders (from 12.5% to 16.0% in engineering, and from 15.6% to 21.2% in CS). However, the share of bachelor’s degrees awarded to Black/African Americans has declined in engineering (from 5.6% to 4.2%) and remained flat (at about 10%) in CS, and the share awarded to American Indians/Alaska Natives has declined in engineering (from 0.6% to 0.4%) and remained flat (at 0.5%) in CS (NSB, 2018).

Many have argued that these issues of underrepresentation could, at least in part, be addressed through more widespread adoption of RBIS. Tobias (1990) and Felder (1993) note that introductory science courses are often responsible for driving off many students who have an initial intention and the ability to earn science degrees but who instead switch to nonscientific fields (i.e., students in the so-called second tier). Women and students from underrepresented minorities are over-represented in this population. Felder (1993) summarized several teaching practices that can inhibit the departure of these students from engineering, many of which comprise RBIS.

Creating physical environments that are inclusive for all genders (Cheryan, Plaut, Davies, & Steele, 2009) and explicitly taking responsibility to promote gender equity in the classroom, rather than exhibiting a gender-blindness approach (Blair, Miller, Ong, & Saxtavker, 2017), have also been postulated as strategies to establish a more inclusive environment among women in STEM. And for other students from groups underrepresented in STEM, several strategies have been suggested: providing more role models and highlighting the important contributions of scientists from underrepresented groups (Beasley & Fischer, 2012), and acknowledging and confronting implicit biases as a way to mind the privilege gap between our students and ourselves when developing our courses and to mitigate stereotype threat in our classrooms (Killpack & Melón, 2016). One engineering department that is employing these strategies is Civil Engineering at Rowan University (Farrell, et al., 2017). A recent e-newsletter from the Executive Director of ASEE announced an ASEE partnership with the CIRTL network to develop materials for “engineering audiences (across the diversity of colleges and universities with engineering and engineering technology programs) focused on (a) research mentorship; (b) inclusive classroom instruction; and (c) undergraduate advisement” (N. Fortenberry, Norman’s Notes, March 26, 2018).

There has been little research on the experience of lesbian, gay, bisexual, transgender, and queer (LGBTQ) individuals in engineering and CS (Cech, Waidzunas, & Farrell, 2017). Since 2014, the American Society of Engineering Education has offered
SafeZone Ally Training to support engineering and CS instructors as they promote LGBTQ equality in STEM (Chavela, Farrell, & Longo, 2016) (https://lgbtq.asee.org/ally-training/). These workshops establish a virtual community of practice among LGBTQ-affirming faculty and supports them with strategies to foster an inclusive environment.

LOOKING TO THE FUTURE

There are many challenges to measuring the extent of RBIS use in engineering and CS. In particular, relying on a single type of measure is likely to result in a biased and incomplete picture of what is happening in STEM classrooms (AAAS, 2013). Further, describing and communicating individual or types of RBIS is an additional significant challenge, even within fields like engineering or CS (Nguyen, et al., 2017b), let alone across STEM fields. We note that engineering is broader than other STEM disciplines in the sense that engineers identify primarily with a particular sub-discipline (e.g., electrical engineering or mechanical engineering), and there may be 10-15 such departments on a single large campus. Thus, the dynamics of diffusion of RBIS are complex in engineering.

Change initiatives in engineering and CS have recently moved in the direction of cultural transformation, guided by a change strategy organizer developed by Henderson, Beach, & Finkelstein (2011) and used in nearly every article in the Journal of Engineering Education special issue on “The Complexities of Transforming Engineering Higher Education” (Vol 103, No. 2). New change initiatives are supported substantially through NSF’s Revolutionizing Engineering and Computer Science Departments (RED) funding program (https://academicchange.org/resources/). The first awards were made in 2015, and publications at this point are primarily limited to press releases about the awards and the departments’ revolutionary plans. Although these sites most certainly have evaluation plans in place, limited information is available about them and no data have been published as of this writing.

An additional future direction for supporting the adoption of RBIS worth mentioning is the systematic literature review (SLR). SLRs are a stand-alone research methodology to address a crafted set of questions by synthesizing primary (and other) studies. They can encourage more instructors to use RBIS by compiling compelling evidence for particular practices, demonstrate gaps in recent work, and highlight areas where a concept is accepted as true, but where little evidence exists to support it (Petticrew & Roberts, 2006). The term SLR refers to an evolving collection of synthesizing methodologies which draw from a wide variety of quantitative (e.g., statistical meta-analysis or network meta-analysis; Tricco, et al. 2011), qualitative (e.g., meta-ethnography or content analysis), and mixed method approaches (Barnett-Page & Thomas, 2009; Gough, Oliver, & Thomas, 2012; Noblit & Hare, 1988), and from studies that have used combinations of these approaches (Thomas & Harden, 2008). Unlike narrative reviews, the criteria for identifying and selecting sources in a SLR are explicitly described in a methods section. Not including this information leaves narrative reviews open to risks of bias and incomplete coverage (Cook & West, 2012).

To be clear, SLRs are not well suited to make generalizable claims about use of RBIS, but they can provide a powerful, evidence-based and often quantitative method to support RBIS use in engineering, CS, and other STEM fields. As a lever for change, SLRs address concerns about the relevance and credibility of evidence that could support STEM
instructor adoption of RBIS. One of the most well-known SLRs related to RBIS in STEM is a study by Freeman and colleagues (2014) that weighed the evidence in support of active learning. Since there have been several SLRs supporting the efficacy of RBIS, our own current research focuses on an SLR of student affective response to RBIS. Our emerging results indicate that (1) there are specific strategies instructors can use (and others they should avoid) to minimize negative student reactions to RBIS, and (2) there are high-quality examples of prior studies which can guide instructors in investigating their own students’ reactions to RBIS.

CONCLUSION

In sum:

▲ Some prior survey studies give us a sense of RBIS adoption rates in engineering and CS.

▲ Our recent research has focused on reasons why engineering and computing instructors do not continue their use of RBIS.

▲ Change interventions are currently focused on long-term, local support of instructors through cultural change and faculty learning communities.

▲ Some attention is being paid to diversity, often in the form of training for instructors, not necessarily in the form of cultural change or faculty learning communities explicitly focused on diversity issues.

We hope that ongoing work and lively discussion with colleagues in engineering, CS, and other STEM disciplines will promote even more widespread adoption of RBIS across the fields.

REFERENCES

American Association for the Advancement of Science (AAAS) (2013). Describing and measuring undergraduate STEM teaching practices. Washington, DC: AAAS.


Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Medical Research Methodology, 8(45), 1-10.


Acknowledgments: Thanks to dozens of colleagues who have shared ideas and conversations about geoscience education that have contributed to this paper. Thanks especially to Heather Macdonald, Cathryn Manduca, David Mogk, and Barbara Tewksbury for suggestions regarding the discussion of the early years of geoscience education, and for their leadership and vision in helping to blaze the path that we follow today.

INTRODUCTION

Some individuals and groups sought to introduce novel content and alternative pedagogies to the geosciences since the mid-1960s (Tewksbury et al., 2013), but the first steps toward community-wide instructional reform were not taken until approximately two decades ago. A 1996 workshop brought together several dozen geoscience instructors, administrators and agency representatives to consider reforms in geoscience education (Ireton et al., 1997). Echoing conclusions from recent STEM publications (NSF, 1996; NRC 1996), the workshop report (and others that would soon follow) issued a call for systematic research on geoscience education and recommended that geoscience instructors should introduce research findings on student learning into their courses (Ireton et al., 1997; Manduca et al., 2002).

Initial efforts to share information about instructional reforms in the geosciences included a Distinguished Speaker Program organized by the National Association of Geoscience Teachers (NAGT), occasional workshops aimed at changing the way instructors taught specific courses (e.g., Teaching Mineralogy) or competencies (e.g., Teaching Quantitative Skills), or more broadly applicable workshops focusing on...
course design or career development (see Manduca, 2008; Tewksbury et al., 2013). In the late 1990s and early 2000s, several programs supported building the capacity for instructional reform in the geosciences through a combination of workshops and/or online materials. The Digital Library of Earth System Education (DLESE)\(^1\) represented an early community-wide effort to provide a venue for college and K-12 teachers to locate educational resources. On the Cutting Edge\(^2\) (CE) created professional development programs, the Teach the Earth\(^3\) portal provided access to online resources related to teaching and learning, and programs such as Starting Point\(^4\) and Pedagogies in Action\(^5\) focused on describing research-based instructional strategies and explaining how they could be applied in geoscience courses. These online sites often became the repositories\(^6\) for resources created by collaborators and geoscientists outside these specific programs. Web materials cataloging the workshop experiences and curating instructional resources were created to enable on-line learning by a broader community of practice representing instructors who could not attend workshops or who were not involved in material development. Links to key programs and projects are included in the footnotes of this document.

There was a corresponding increase in the number of half-day or full-day sessions devoted to geoscience education at the meetings of professional societies such as the Geological Society of America (GSA) or American Geophysical Union (AGU). For example, in the mid-1990s, a handful of sessions on geoscience education were held.

---

\(^1\) Digital Library for Earth System Education, http://www.dlese.org
\(^2\) On the Cutting Edge, https://serc.carleton.edu/NAGTWorkshops/index.html
\(^3\) Teach the Earth, https://serc.carleton.edu/teachearth/index.html
\(^4\) Starting Point, https://serc.carleton.edu/introgeo/index.html
\(^6\) SERC Projects and Collaborations, https://serc.carleton.edu/serc/site_guides/projects.html
at the annual GSA meeting. By the time the society met in Seattle in 2003, there were 17 sessions featuring geoscience education themes. These sessions were increasingly populated by faculty who were reporting results from the early stages of geoscience education research (GER). This research often featured the adaptation of strategies that had been shown to be successful in other STEM disciplines. For example, Libarkin and Anderson (2005) discussed data from the deployment of a Geoscience Concept Inventory (GCI), McConnell et al. (2006) created and shared a set of geoscience ConcepTests®, and Kortz, Smay and McMurray (2008) assessed the application of lecture tutorials in introductory courses. These approaches were motivated by examples from physics education research (Halloun & Hestenes, 1985; Mazur, 1997; Shaffer & McDermott, 1992, respectively).

Around this time, GSA published its first edited volume on geoscience education, *Earth and Mind: How Geologists Think and Learn about the Earth* (Manduca & Mogk, 2006; Manduca, 2008), that focused on themes such as spatial reasoning, temporal thinking, complex systems, and learning in the field (see also Kastens & Manduca, 2012). The major geoscience professional organizations (GSA, AGU, NAGT and the American Geosciences Institute [AGI]) also collaborated to support the Building Strong Geoscience Departments® (BGSD) project that helped guide faculty and departments through critical disciplinary and institutional changes.

Change efforts were not limited to the professional organizations and early GER faculty. Numerous federally funded geoscience research projects included an educational component to broaden the impact of their work. Programs such as Earthscope®, IRIS®, UNAVCO®, and Margins/GeoPRISM® included professional development opportunities for faculty and created educational resources that could be readily incorporated into college courses (and often K-12 settings). With time, the NAGT’s Distinguished Speaker Program and the Building Strong Geoscience Departments workshop program were replaced by the more broadly defined Traveling Workshops Program® (Manduca et al., 2005). The cumulative result of these activities was an increased interest in geoscience education and its applications. When the annual GSA meeting returned to Seattle in 2017, it featured more than 40 geoscience education sessions.

The initiatives outlined above were operating in parallel with On the Cutting Edge (CE), which has remained a significant national-scale professional development program in the geosciences. CE developed models for a range of workshop formats and lengths and offered 143 workshops between 2002 and 2016. The CE program was managed by a team of PIs but many more were involved; over the life of the project, hundreds of geoscience faculty were recruited to plan and run workshops and to contribute resources to the website. CE sought to build a geoscience education community of practice through (1) networking at workshops; (2) communication via email lists; and, (3) creation of resource collections®. Tewksbury et al. (2013) noted that, in their experience, geoscience instructors “are not interested in turn-key

---

7 Geoscience ConcepTest collection, [https://serc.carleton.edu/introgeo/interactive/conctest.html](https://serc.carleton.edu/introgeo/interactive/conctest.html)
8 Building Strong Geoscience Departments, [https://serc.carleton.edu/NAGTWorkshops/departments/](https://serc.carleton.edu/NAGTWorkshops/departments/)
9 Earthscope Education, [http://www.earthscope.org/education](http://www.earthscope.org/education)
10 IRIS Education, [https://www.iris.edu/hq/programs/epo](https://www.iris.edu/hq/programs/epo)
11 UNAVCO Education, [https://www.unavco.org/education/education.html](https://www.unavco.org/education/education.html)
12 MARGINS research in the classroom, [https://serc.carleton.edu/margins/index.html](https://serc.carleton.edu/margins/index.html)
13 Traveling Workshops Program, [https://www.nagt.org/nagt/profdev/twp/](https://www.nagt.org/nagt/profdev/twp/)
14 On the Cutting Edge Program Impact, [https://serc.carleton.edu/NAGTWorkshops/about/project_impact.html](https://serc.carleton.edu/NAGTWorkshops/about/project_impact.html)
curriculum materials, preferring instead either to develop their own resources or to adapt resources from a variety of sources in order to meet the needs of their own students, departments, and institutions” (p.198.) Consequently, CE was not focused on promoting a specific set of instructional practices or curricula but sought to build a diverse skill set among geoscience instructors that would result in (1) Increased instructor knowledge of effective teaching practices; (2) Creation of effective online resources for distributing workshop content and appropriate instructional materials; and, (3) Building a cohort of educators to disseminate key aspects of course and curriculum design and reformed instruction (Tewksbury et al., 2013).

Workshops were designed to cater to a range of needs and interests among geoscience instructors. Workshop types included:

- **Emerging theme** workshops to introduce new geoscience research or a new aspect of pedagogy (e.g., The Role of Metacognition in Teaching Geoscience; Teaching Geoscience Online).

- **Teaching X** workshops brought together instructors who were teaching a core geoscience topic that was typically part of the geoscience major (e.g., Teaching Structural Geology, Teaching Mineralogy).

- **Recurring annual workshops** supported faculty at various career stages (e.g., Preparing for an Academic Career; Building Strong Geoscience Departments) and **large workshops** disseminated best practices in teaching, often emphasizing introductory courses.

CE also offered webinars, webinar series, and journal clubs. Further, there was collaboration between geoscience researchers and the CE program to host workshops at professional meetings on teaching particular scientific topics. While CE has produced an extensive suite of professional development opportunities, we should keep in mind that geoscience faculty were also likely to have been exposed to similar programs through their institutions, professional organizations, and other groups that included instructional reform as part of their mission (e.g., PKAL, AAC&U).

This chapter reviews recent publications and reports to examine how successful the geosciences have been in advancing research-based reforms in undergraduate instruction. We describe quantitative and qualitative data and infer that instructional reform is widely practiced in the discipline, largely as a consequence of broad participation in many of the programs mentioned above.

**RESEARCH-BASED REFORMS**

The CE program collected data about teaching practices in undergraduate geoscience classes at US institutions using multiple self-report surveys, a teaching observation project, and detailed interviews with geoscience instructors. These data can be supplemented and supported with similar information from other sources.

**NATIONAL SURVEYS AND OBSERVATIONS**

The CE project conducted three national surveys (2004, 2009, 2012) of geoscience instructors at US higher education institutions (Manduca et al., 2017). While these data were collected by the CE program, it did not explicitly target CE participants but rather requested information from as broad a population of geoscience instructors as possible. Depending on the iteration, surveys were sent to between 5700 and 7813
Geoscience education researchers collected direct measurements of instructional practices from more than 200 geoscience instructors in a range of classrooms and institutions (Teasdale et al., 2017; see also, Budd et al., 2013). Classroom observations in the geosciences have relied on the Reformed Teaching Observation Protocol (RTOP; Sawada et al., 2002) to measure teaching practices. The RTOP instrument describes teaching processes on five subscales of five items each. The resulting 25 items are each scored on a scale of 0-4 resulting in a possible total score of 0-100 (Sawada et al., 2002).

A pair of observers had previously created a scoring rubric to guide use of the RTOP to describe the teaching practices of 26 instructors in 66 introductory geoscience classes at a variety of institution types (Budd et al., 2013). Subsequently, a team of trained classroom observers, sponsored by the CE Classroom Observation Project16 (COP), used a modified version of the RTOP rubric to complete more than 200 observations in a wide range of geoscience classrooms across the United States (Teasdale et al., 2017). These observations were conducted within a few years of the third national survey (2011-2014). The COP study also examined relationships between teaching practices and variables such as class size, class level, and institution type and matched RTOP scores with observer comments to characterize the types of teaching practices that were characteristic of each RTOP category (Teasdale et al., 2017). Participating instructors also responded to a subset of questions on instructional practices selected from the CE survey.

**Interview Data**

The surveys and classroom observations reported above included a variety of instructors, some of whom had participated in CE programming and many who had not. The interview data reported here represents 120 instructors who had participated in CE programming, and the interviews sought to identify which aspects of professional development supported instructional change. Nearly 60% of the interview participants had also completed the CE survey, allowing individual responses to be compared. The interviews provided insights into how CE resources influenced instructional reforms, attitudes and beliefs about teaching, and motivation and enthusiasm for change (Manduca et al., 2017).

**Limitations**

The combination of surveys, direct observations and participant interviews provides a robust set of data to analyze for patterns and trends. However, there are limitations to the information we can derive.
from such analysis. It is estimated that there may be approximately 10,000 geoscience instructors in higher education in the US (Manduca et al., 2017); consequently the “true” sample size may be closer to 20-30%. The study respondents may not be representative of the overall population of geoscience instructors. Respondents represent different populations for each survey, with just 454 instructors completing all three surveys (1007 completed both 2009 and 2012 surveys). Later surveys were more likely to feature responses from instructors at two-year colleges, as the researchers were able to administer the survey to a more representative population of faculty (Manduca et al., 2017). The data were more likely to represent instructors in geology programs, with fewer responses from atmospheric science or marine science departments. Other potential limitations are discussed in Manduca et al. (2017).

**WHAT DO THE DATA TELL US?**

Over the last decade, research-based instructional strategies have become a more common feature of undergraduate geoscience courses. Reforms can be broadly defined as research-based instruction without emphasizing a particular curriculum or pedagogical approach. These reforms are more likely to be present in courses taught by instructors who have been involved in professional development.

### Survey Results

The proportion of geoscience faculty who self-identify as having incorporated research-based instructional strategies in their courses has climbed from about a third in the initial survey in 2004, to just over half of survey respondents in the third version of the survey in 2012 (Manduca et al., 2017). A comparison of two survey questions revealed a statistically significant upward trend in the use of research-based instructional methods over time (Manduca et al., 2017, Table 5.1).

The first survey question asked instructors to estimate what proportion of each class students were actively engaged in activities, questions and discussion. In the 2012 survey, over half of participants (51%) indicated that these methods accounted for more than 20% of their class time. The next question asked instructors to estimate how much of their class time was dedicated to several different instructional strategies including (a) traditional lecture; (b) lecture with demonstration; (c) lecture with questions answered by individual students; (d) lecture with questions answered simultaneously by the whole class; e) small group discussions or think-pair-share; f) whole-class discussions; and, g) in-class exercises (Table 5.1). In the 2012 survey, 57% of respondents indicated that they used one of the latter three strategies (e, f, g) frequently (weekly or in nearly every class). Further, results from these two questions show a moderate

<table>
<thead>
<tr>
<th>Survey question/year</th>
<th>2004</th>
<th>2009</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the “lecture portion” of your _________ course, please estimate the percentage of class time spent on student activities, questions, and discussion (proportion of responses that were &gt;20%)</td>
<td>34% (n=1566)</td>
<td>38% (n=2030)</td>
<td>51% (n=1766)</td>
</tr>
<tr>
<td>In the “lecture portion” of your _________ course, please indicate how frequently you used small group discussion, whole-class discussion, or in-class exercises with or without the use of any other methods (proportion of responses for “weekly” or “nearly every class”)</td>
<td>42% (n=1568)</td>
<td>47% (n=2017)</td>
<td>57% (n=1741)</td>
</tr>
</tbody>
</table>
correlation ($r(1711) = 0.46, p < 0.001$) (Manduca et al., 2017).

Manduca et al. (2017) used a cluster analysis to identify three distinct groups of faculty in the survey data on the basis of their primary professional responsibilities: education-focused faculty (20% of 2012 survey population), disciplinary research-focused faculty (37%), and teaching faculty (43%). Education-focused faculty were identified in all types of institutions and reported significant activity related to improving their own teaching or the teaching of others. Teaching faculty reported lower levels of activity in geoscience research than education- and research-focused faculty and were less likely to have attended talks or workshops related to improving teaching than were their education-focused peers. Instructors in the education-focused group were the most likely to attend teaching-related talks or enroll in workshops at professional meetings. Perhaps unsurprisingly, this subgroup of instructors was the most likely to report the use of active-learning teaching strategies. However, all three groups contained a substantial proportion of faculty whose survey responses suggested that they employed research-based instructional strategies (Table 5.2; Manduca et al., 2017).

Research-focused faculty typically reported spending about half as much time teaching as members of the other subgroups. However, even among these faculty, the proportion who had attended three or more education talks (23% vs. 37%) and one or more teaching workshop (25% vs. 38%) increased between administration of the last two surveys (2009, 2012). A similar trend was observed for teaching faculty (28% vs. 43% for education talks; 36% vs. 52% for workshops; Manduca et al., 2017).

### Observation Results

Both the Classroom Observation Project (Teasdale et al., 2017) and the earlier study by Budd et al. (2013) characterized teaching practices in geoscience classrooms on the basis of total RTOP scores, as teacher-centered (≤30), transitional (31-49), or student-centered (≥50), and found similar ranges (13-89 vs. 18-87) and mean scores (39.6 vs. 41.5). The Classroom Observation Project found that the 204 observed faculty could be subdivided as 30% teacher-centered, 45% transitional, and 25% student-centered (Teasdale et al., 2017).

Classroom observations revealed that there was no single teaching strategy that predicted the general use of research-based instructional practices. Teasdale et al. (2017) noted that there seven of the 25 total RTOP items were most likely to be predictive of RTOP score. Five of these items came from two subscales of the RTOP instrument (Student-student Interactions and Student-instructor Interactions). The key categories and items are as follows (Teasdale et al., 2017):

#### Table 5.2. Survey responses for “active learning” questions, 2012 survey. (Data from Table S9, Manduca et al., 2017)

<table>
<thead>
<tr>
<th>Survey question/faculty subgroup</th>
<th>Education-focused faculty</th>
<th>Research-focused faculty</th>
<th>Teaching faculty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of faculty who spent &gt;20% of the “lecture portion” of their course, on student activities, questions, and discussion</td>
<td>67%</td>
<td>44%</td>
<td>49%</td>
</tr>
<tr>
<td>Proportion of faculty indicating they used small group discussion, whole-class discussion, or in-class exercises in the “lecture portion” of their course</td>
<td>72%</td>
<td>51%</td>
<td>55%</td>
</tr>
</tbody>
</table>

**Category 1: Lesson Design**

**Item 2: The lesson was designed to...**
engage students as members of a learning community.

▲ **CATEGORY 3: Procedural Knowledge**

**Item 13:** Students were actively engaged in a thought-provoking activity that often involved the critical assessment of procedures.

▲ **CATEGORY 4: Student-Student Interaction**

**Item 16:** Students were involved in the communication of their ideas to others using a variety of means and media.

**Item 18:** There was a high proportion of student talk and a significant amount of it occurred between and among students.

**Item 19:** Student questions and comments often determined the focus and direction of classroom discourse.

▲ **CATEGORY 5: Student-Instructor Interaction**

**Item 24:** The teacher acted as a resource person, working to support and enhance student investigations.

**Item 25:** The metaphor “teacher as listener” was very characteristic of this classroom.

**What was happening in classes?**

Aside from RTOP scores, observer comments on instructional practices reflected the range of application and frequency of use of these seven items and others (Figure 5.1). For example, considering Item 18 that measured the level of student-student talk during a class, observers noted that most teacher-centered classes were characterized by lecture and passive students who did not interact with each other. Some instructors would ask questions that elicited “shout-out” responses and had students who occasionally asked questions or volunteered ideas, but over 90% of the observation data reported no student-student talk during the observed lesson.

Instructors in transitional classrooms were more likely to encourage peer discussion and the use of ConcepTest-type questions (see Mazur, 1997; McConnell et al., 2006), but more than 80% of the observation data for the transitional classrooms reported student-student talk for less than 10% of the time. In the student-centered classrooms, observers most frequently noted increased levels of student interactions and discussions, students asking questions or sharing ideas, and the use of activities that required students to interact with data, with nearly 60% of the classes recording over 25% student-student talk during the observation (Figure 5.1). The amount of time devoted to lecture decreased from teacher-centered to student-centered classrooms, while there was a corresponding increase in the frequency of the assessment of student learning (Teasdale et al., 2017).

**Figure 5.1.** Commonly observed attributes of geoscience classrooms as determined by qualitative analysis of observer comments (modeled after Table 4, Teasdale et al., 2017).
Instructors incorporating research-based instructional strategies were likely to use strategies with high instructional utility, in that they often required minimal preparation, were relatively straightforward for students to use, and could be used frequently in classes of any size (see discussion in McConnell et al., 2017). However, some strategies are better backed than others by empirical evidence (e.g., several studies, consistent evidence of learning gains in multiple settings) to validate their use as aids in promoting student learning. For example, observations of classroom teaching describe examples of peer instruction and think-pair-share. Both strategies have high utility, but peer instruction is supported by a robust collection of research studies (e.g., Crouch & Mazur, 2001; Smith et al., 2009) to support its learning efficacy, while research studies mentioning think-pair-share usually include it among a suite of strategies (e.g., Yuretich et al., 2001) and rarely single out its impact on student learning.

The instructional strategies employed in geoscience classrooms align with the characterization by Stains et al. (2018) of three general types of instruction observed in national STEM classrooms. Their category ‘didactic instruction’ matches closely with the traditional category of geoscience instruction as defined by RTOP scores of ≤30; their ‘interactive lecture’ style seems to align with the transitional category of geoscience classrooms; and their ‘student-centered’ instruction is a general match for the same category defined by Teasdale et al. (2017).

Observed faculty completed a subset of survey questions similar to the national survey described earlier. Their survey results can be categorized on the basis of the three categories of instructional practices. A majority of instructors in student-centered classrooms reported frequently using small group discussions or think-pair-share activities and in-class exercises. They also reported posing questions to individual students and to the entire class. Most
of these instructors also reported using traditional lecture every class period. Less than half of instructors in transitional classrooms reported frequent use of small group discussion or think-pair-share activities and only a third noted frequent use of in-class exercises. These instructors reported being more likely to use traditional lecture and were slightly less likely to ask individual or whole class questions. Almost all of the instructors in teacher-centered classrooms stated that they used traditional lecture during every class period. Only 10% used small group discussions or in-class exercises weekly, and approximately a third reported daily use of questions to individual students and to the whole class (Teasdale et al., 2017).

Results on the Basis of Class, Institutional, or Instructor Characteristics.
While a range of instructional strategies are being deployed in geoscience classrooms, there does not appear to be any relationship with the use of these strategies and a variety of institutional or demographic characteristics. There was no difference in RTOP scores on the basis of the type of institution (Research/Doctoral, Masters, Baccalaureate, Associate) where the courses were taught, the academic level of the course (introductory vs. majors), the size of the class, or gender of the instructor (Teasdale et al., 2017).

There was a significant difference in RTOP scores on the basis of professional development experience. CE workshop participants were more likely to have taught using research-based instructional strategies, and therefore earned significantly higher RTOP scores than instructors who had not participated in CE programs or who only used the online resources (i.e., had not attended a workshop). The mean RTOP score for instructors who had participated in a CE workshop and used the website was 48.2; while the equivalent score for instructors who had not attended a workshop or used the website was 33.1 (see Figure 4, Table S16 in Manduca et al., 2017). Instructors who attended one or more CE workshops and who also made use of the resources available through the related website were 1.5 times more likely to spend more than 20% of class time on activities, questions, or discussion. Further, they were 1.2 times more likely to report teaching using an active learning teaching strategy such as small group discussion, whole-class discussion, or in-class exercises, than instructors who had neither used the website nor attended a CE workshop (Manduca et al., 2017).

Interview Results
Analysis of interview transcripts on the impact of CE revealed that nearly three quarters (72%) of participants noted specific changes to their instructional practice as a result of
workshop participation and using related online resources (Manduca et al., 2017). Many of these instructors also noted a pedagogical shift in which changes in their practice were accompanied by changes in their attitudes or beliefs about teaching. Instructors who participated in CE stated that they made a variety of instructional changes and did not rely on a few specific or prescribed instructional models. Further, results suggest progressive changes as instruction evolved and incorporated multiple components of research (Manduca et al., 2017).

Effective professional development experiences can help instructors adopt research-based instructional strategies and can also result in a permanent realignment of the instructor’s teaching beliefs. University geoscience instructors were guided in the development of reformed instructional materials for introductory college geoscience courses as part of the InTeGrate (Interdisciplinary Teaching about the Earth for a Sustainable Future) project. Pelch and McConnell (2016) used the Teacher Belief Interview (TBI) (Luft & Roehrig, 2007) and participants’ reflections on their experience to characterize pedagogical beliefs at different stages of the professional development process. The TBI is a semi-structured instrument consisting of seven questions that help researchers characterize respondents’ views on teaching and learning across five categories (traditional, instructive, transitional, responsive and reformed) from traditional to reform-based instruction. There was a significant change in the majority of InTeGrate participants’ TBI scores, which improved toward more reform-based pedagogical beliefs ($t(20) = 3.43, p = 0.003, d = 0.70$). Instructors who began with the most traditional pedagogical beliefs showed the greatest gains. The most significant changes occurred in areas strongly influenced by situational classroom factors. Three TBI questions in particular accounted for the greatest proportion of the change in score: How do you know when to move on to a new topic in class? How do you decide what to teach and what not to teach? How do you know when your students understand a concept in class? (Pelch & McConnell, 2016).

**WHAT WORKED AND WHY?**

Henderson et al. (2010) defined four principal types of change in STEM education: (i) Disseminating Curriculum and Pedagogy; (ii) Developing Reflective Teachers; (iii) Enacting Policy; and (iv) Developing Shared Vision. Investigations into change within the geosciences suggest that aspects of three of these categories (i, ii, and iv) are related to instructor-generated change, as opposed to change driven by top-down processes.

**Disseminating Curriculum and Pedagogy**

While there has been less emphasis in the geosciences than some other STEM disciplines on developing standalone curricula, the geoscience community
has created an extensive professional development program\textsuperscript{2} and supporting website\textsuperscript{3} (>5000 pages, topical collections, community-contributed teaching activities) that contains a variety of peer-reviewed workshop products and other teaching and learning resources (Manduca et al., 2010). Approximately a quarter of geoscience instructors have participated in CE workshops and many report changes in their pedagogical beliefs as a result. Further, the robust use of the website resources (more than a million unique visits per year; Manduca et al., 2017) suggests that the instructional resources are valued by others who have not participated in CE workshops.

**Developing Reflective Teachers**

Leaders report that professional development in the geosciences has created a culture of sharing and communal improvement in support of undergraduate geoscience teaching (Manduca et al., 2010; Kastens & Manduca, 2017). Professional development design principles for CE emphasized introducing ideas from cognitive science and education research, engaging participants in reflecting on applications to their teaching, and learning from peers (Manduca, 2017). These strategies may incorporate some aspects of established principles for effective professional development (NRC, 2012), such as feedback on instructional practice and a focus on changing conceptions about teaching and learning. For example, learning from peers may include receiving explicit or indirect feedback on instructional practice; reflecting on applications to one’s own teaching may contribute to changing conceptions about teaching and learning (NRC, 2012; Pelch & McConnell, 2016).

**Developing Shared Vision**

Kastens and Manduca (2017) suggested that the combination of individual learning, supportive colleagues, and group accomplishments have been important in building a community of practice within geoscience education. Individuals learn from trusted colleagues teaching in similar contexts or from outside experts sharing new ideas. This process ensures that participants leave workshops with ideas that they can readily apply in their own classes. Peers are considered a trusted source of information and provide practical advice on the basis of their own experiences (Manduca, 2017). Professional development programming provides opportunities for instructors to interact with supportive colleagues who share a common desire to improve teaching and learning. This can be especially significant for instructors who are the only geoscience faculty member on their campus or the only one seeking to promote educational reform.

Interview feedback reveals that one aspect of the CE program that proved pivotal in encouraging change for many instructors (73\%) was the opportunity to share ideas with peers and to learn from the experiences of colleagues (Manduca et al., 2017). Finally, professional development in the geosciences often occurs around group accomplishments of a specific task such as the structured review of online resources or the creation of instructional resources (e.g., McConnell et al., 2013). Both the individual participant and the whole population of geoscience educators benefit from this model of community practice (Kastens & Manduca, 2017).
ONGOING REFORM EFFORTS

The early reforms in geoscience education were largely concerned with the introduction of research-based instruction and adapting instructional practices to better teach critical geoscience concepts. The next generation of reforms reflects the growing diversity of leadership in geoscience education research and focuses on a range of challenges such as the influence of the affective domain on student learning, the adoption of instructional materials that better link geoscience and society, and strategies to foster inclusion and diversity within the discipline.

CREATING A FORMAL GEOSCIENCE EDUCATION RESEARCH COMMUNITY

Beyond participation in CE programming, there have been other efforts to build a stronger geoscience education community among faculty at higher education institutions and, thus, provide a structure to promote the adoption of research-based instructional strategies. While research on geoscience education has been occurring for at least two decades, a Geoscience Education Research (GER) division (Lukes et al., 2015) was only recently established within the National Association of Geoscience Teachers (NAGT). The GER division has created a GER toolbox that shares instruments and tools and guides researchers through project design.

In addition, the GER division has taken the lead in a community-wide “Grand Challenges” conversation that seeks to identify priorities for new geoscience education research that can positively impact undergraduate teaching and learning in the geosciences (St. John et al., 2017). Forty GER experts contributed to the development of a framework of 10 potential research themes (e.g., Access and Success of Underrepresented Groups in the Geosciences) and a series of related Grand Challenges. For example, two grand challenges associated with the Access and Success topic are: (1) How can we recognize and support the individual identities and personal pathways of students as they are attracted to and thrive in the geosciences? (2) How can the geoscience community capitalize on evidence from different scale efforts to broaden participation? Each challenge is presented with potential strategies to meet the challenge. The project seeks to set a series of research goals for GER that can be achieved within a decade. Feedback from a combination of webinars, discussion forums, town hall meetings, surveys and peer reviews was solicited to produce a final document (St. John, 2018).

EARTH EDUCATORS’ RENDEZVOUS (ANNUAL MEETING)

The success of the CE program and other initiatives highlights the need for sustainable professional development experiences for geoscience instructors. To meet this need, the Earth Educators’ Rendezvous was conceived as a stand-alone, five-day professional development event. The Rendezvous attracts more than 300 GER scholars, geoscience educators, and colleagues each summer, offering a mix of programming including multi-day workshops, mini-workshops, research presentations, round table discussions, working groups, and plenary sessions. Rendezvous financial support largely comes from the registration fees of participants along with support from some researchers or agencies who use the Rendezvous as a venue for their funded workshops.

19 Earth Educators’ Rendezvous, https://serc.carleton.edu/earth_rendezvous/
Building Effective Resources for Geoscience Teaching

Over the last 20 years there has not only been a shift in how the geosciences are taught, but also in what is taught under the banner of geosciences. In many introductory courses there is now less emphasis on traditional technical content, basic facts and concepts of how Earth works, and more attention to how the geosciences impact the lives of individuals and society at large (Bralower et al., 2008). The recent InTeGrate project\(^{20}\) (Interdisciplinary Teaching about Earth for a Sustainable Future) may be considered a representation of these types of resources that are both better aligned with research-based design practices and conceived to emphasize the link between the geoscience and societal issues (e.g., environmental justice). The InTeGrate project brought more than a hundred geoscience instructors together to develop new instructional materials that were designed to include research-based instructional strategies. More than 30 modules or courses were generated around an aligned set of learning objectives, formative and summative assessments, along with classroom activities that emphasized student-student and student-instructor interactions (McConnell et al., 2013).

InTeGrate activities were created to reduce the amount of time the instructor spends lecturing and increase the proportion of class time students spend working collaboratively on problems and participating in discussions. Each activity within a module or course was developed employing backward design (Wiggins & McTighe, 2005). This style of course content creation centers on first establishing the measurable tasks instructors want students to know (learning objectives), then designing the activities students will perform to meet those objectives, and finally on creating formative and summative assessments to measure student learning. It is unlikely that such a broad-ranging initiative could have been successful without the support of a community of educators who were at least partially familiar with the principles of research-based instruction and who recognized the value in expanding the collection of resources that were available. Many of the resource developers and team leaders within InTeGrate developed the skills and experience to take on these roles through their prior experiences in CE and other professional development programs.

Educating Leadership About Geoscience Education

The summit on the Future of Undergraduate Geoscience Education (Mosher et al., 2014) brought together academics and geoscience employers to consider (a) the knowledge, competencies and skills that future geoscientists would need to build successful careers; (b) the best ways to use teaching and technology to enhance learning; and (c) how to increase participation of underrepresented groups in the geosciences. The summit, and the meetings that followed, resulted in a series of recommendations for improving undergraduate programs and curricula. Among the Summit’s recommendations was that research-based instructional strategies should be incorporated more widely in geoscience classrooms:

Surveys reveal that increasing proportions of geoscience faculty are implementing such active-learning pedagogies over time, but the penetration of such approaches across departments, and even faculty awareness of them is still limited. Hence, our community’s primary pedagogical challenges lie in encouraging the wider adoption of these practices and in characterizing their educational impacts and benefits (emphasis in original)\(^{21}\) (Mosher et al., 2014, p.4).
The summit organizers plan to produce a Vision and Change document in the near future to help direct departments and programs seeking to adopt these best practices.

**IMPROVING DIVERSITY AND INCLUSION IN THE GEOSCIENCES**

Although numbers of students from African American, Hispanic, and other minority groups in geoscience programs have increased, the geosciences still fall below national benchmarks for participation for all minority populations (NSB, 2018). The National Science Foundation’s Opportunities for Enhancing Diversity in the Geosciences (OEDG, 2002-2013) was aimed at increasing the diversity of geoscience programs (Callahan et al., 2017; Wolfe & Riggs, 2017). The proportion of undergraduate natural science degrees awarded to Hispanic and African American students was 10% and 6.7%, respectively, in 2015. In comparison, smaller proportions of geoscience undergraduate degrees were awarded to Hispanic (7.5%) and African American (2.3%) students during the same year (NSB, 2018). However, this represents a near doubling of equivalent values from a decade earlier (3.8%, 1.7%, respectively). A similar pattern is observed at the doctorate level. In 2015, Hispanic students and African American students each accounted for approximately 4% of natural science PhDs. In contrast, of the 827 PhDs awarded in the geosciences, 3.7% went to Hispanic students and 1.3% to African American students. This marked an increase from 2005 (1.9%, 0.5% respectively; see also Sidder, 2017).

Efforts to broaden participation in the geosciences have been ongoing for some time (see Riggs & Alexander, 2007). Wolfe and Riggs (2017) identified approaches that included one or more of the following strategies: (a) effective mentoring (often in association with research opportunities); (b) institutional peer support groups (e.g., learning communities); (c) bridge programs occurring during the transition from high school to college; (d) the incorporation of pedagogies that better support student learning (e.g., place-based learning programs); and (e) programs that address campus climate, culture and support for diversity and inclusion. Programs that have been successful in recruiting and supporting underrepresented students were resource-intensive and often received support from both within and beyond the institution (Wolfe & Riggs, 2017). Wolfe and Riggs (2017) suggest that any efforts to improve diversity and inclusion in undergraduate geoscience programs will require a thorough understanding of existing institutional support structures and a sustained commitment from department faculty and administrators.

**SUMMARY**

Self-report survey data collected at three separate times over an eight-year span can be interpreted to reveal that geoscience instruction at US colleges and universities has experienced a steady change to include
more research-based instructional methods. More than half of survey respondents in the most recently reported survey (2012) identify using active learning strategies in their classes (Manduca et al., 2017). All types of faculty, from those with an interest in geoscience education to those with a focus on basic geoscience research, show an interest in incorporating these strategies into their classes.

These self-reports are supported by classroom observations and instructor interviews. Due to the choice of a specific observation instrument, the classroom observations defined instructional strategies differently from the surveys. Observation results suggest that a quarter of instructors were using extensive active learning and/or research-based instruction, while another 45% used some elements of these strategies (Teasdale et al., 2017). This interpretation is supported by data from a separate study of teaching across multiple STEM disciplines in which Stains et al. (2018) noted that, relative to chance, geology classes featured more student-centered instruction than expected. (Geology is the dominant form of geoscience taught in the US.)

These observations also reveal that there is no difference in the character of instruction on the basis of type of institution, size of class, class content or the gender of the instructor. Further, the nature of this instruction associated with different levels of reform appears to be similar regardless of whether it is described for geoscience courses (Teasdale et al., 2017) or in other STEM disciplines (Stains et al., 2018). In particular, student-centered courses are characterized by peer interactions in small groups that occupy a considerable portion of class time, and often include student-instructor interactions and assessment activities.

Finally, changes in instructor beliefs about instruction appear to be contemporaneous with changes in instructional practices. It remains unclear whether the principal drivers of changing beliefs are external (e.g., information from workshops), the instructor’s personal experience of the changes in practice, or changes to student learning and/or engagement. It is likely that these factors are interconnected (Clarke & Hollingsworth, 2002) and collectively motivate change (Pelch & McConnell, 2016). We anticipate that different instructors are strongly influenced by different drivers.

**FACTORS AND PROCESSES**

Instructional changes are more likely to be observed in classes when the instructor has participated in professional development and made use of related online resources (Manduca et al., 2017). More than 3000 faculty (~30% US instructors), post-docs, and graduate students have participated in one or more CE workshops, and over a quarter (28%) of the participants attended two or more workshops. US geoscience faculty representing more than half the nation’s geoscience departments have participated. CE workshop participants included a higher percentage of women and members of underrepresented groups than in the overall population of geoscience faculty. While it is likely that workshop participants were also engaged in other professional development experiences during this time, there is a clear signal that CE workshop participation is linked to the use of more student-centered pedagogy. In addition, interview transcripts suggest that most CE participants attributed specific changes in their instructional practice to workshop participation and the use of associated online resources (Manduca et al., 2017).
The apparent success of professional development in the geosciences may partially be a function of the fact that there are fewer faculty in geoscience (Table 5.3) than in other STEM disciplines. Thus, the same number of workshops etc. in other disciplines would have had a proportionally smaller impact on instructional reform. Alternatively, the information presented here may be scaled appropriately to determine the approximate number and type of professional development activities that would be sufficient to produce equivalent reforms in other disciplines.

The geosciences have a smaller education research footprint than many other STEM disciplines. For example, Freeman et al. (2014) identified more than 20 studies each in physics, chemistry, biology, and mathematics that supported their contention that active learning in these disciplines increased student performance. Their report included only two studies in geology (Freeman et al., 2014). While the Journal of Geoscience Education has been published since 1951, for much of that time it served as a forum for exchanging teaching ideas and has only become a consistent source for geoscience research articles in the last decade (Piburn et al., 2011). Consequently, to this point, instructional change in the geosciences is unlikely to have been driven by a rich collection of research studies that support its positive impact on student learning. However, this may also have contributed to the development of a community of practice where participants had no association with specific curricular materials and were therefore open to incorporating whatever reforms they believed were best suited to their particular class setting and instructional experience.

The nature of reform in the geosciences has focused more on generally applicable teaching strategies and customizable resources than the development of fully-fledged curricula or a reliance on a specific type of instructional reform. The relative standardization of content in paired introductory courses in some STEM disciplines isn’t matched in the geosciences, where courses are more likely to feature a variety of topical themes and where students rarely take a second course (see Tewksbury et al., 2013).

The future of instructional reform in the geosciences will benefit from a GER focus on a set of goals identified by the community, such as how to more effectively teach specific geoscience content (e.g., climate science) and skills (e.g., quantitative reasoning), how to increase access to the geosciences, and which instructional strategies are most likely to improve student learning. This may require identifying novel ways to encourage participation in professional development (e.g., Earth Educators’ Rendezvous) or developing resources that achieve emerging goals such as making a stronger connection between geoscience and society (e.g., InTeGrate project). Regardless of what individual instructors choose to do, without some departmental support their efforts may come to nothing if changes are not institutionalized so that course and curriculum reforms become established within departmental culture.

Table 5.3. Number of post-secondary teachers (instructors) in STEM disciplines at Colleges, Universities, Professional Schools and Junior Colleges. (Bureau of Labor Statistics data for May 2017, retrieved April 15, 2018.)

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Biology</th>
<th>Chemistry</th>
<th>Engineering</th>
<th>Geosciences</th>
<th>Mathematics</th>
<th>Physics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of positions</td>
<td>49,180</td>
<td>21,080</td>
<td>37,310</td>
<td>10,690</td>
<td>50,630</td>
<td>13,540</td>
</tr>
</tbody>
</table>
REFERENCES


INTRODUCTION

A common narrative around undergraduate mathematics instruction typically involves a research-obsessed, disinterested teacher, standing at the front of a lecture hall, back turned to his students, writing walls of text and computations for a room of disengaged students. Explanations for this instructional paradigm, typically not based on empirical data, are summarized by Weber (2004) as follows:

Kline (1977) argues that because of pressures to publish, mathematicians have neither the time nor the desire to teach advanced mathematics courses well. Thus, in a vicious cycle, these mathematicians resort to teaching their courses in the dry manner in which they were taught. Davis and Hersh (1981) suggest that some professors use DTP [definition, theorem, proof] instruction because they have the desire to appear brilliant; by presenting mathematical theory as a polished product, students will be impressed by these professor’s deductive abilities. They also suggest that other professors are forced to teach straight from the textbook because they do not really understand the subject that they are teaching. Leron and Dubinsky (1995) believe that professors may teach the way they do because these professors believe their instruction will be ineffective no matter what they do; advanced mathematical concepts are simply too difficult for most students to understand in a semester-long course. In short, some conjecture that professors teach using a DTP format because they are disinterested, arrogant, following custom, or insecure about their understanding of the mathematics involved (p. 131).

This stereotypical portrayal of mathematics instruction, and the narrative that mathematics instruction is rooted in habit or apathy, is both unproductive and inaccurate.

Working Group: For a full description of the working group, please see Appendix A. Pictured here from left to right: Shonda Kuiper, Grinnell College; David Webb, University of Colorado Boulder; Rachel Levy, Convener, Harvey Mudd College; Natasha Speer, University of Maine; Estrella Johnson, Reviewer, Virginia Tech; Emily Miller, AAU; and Ann Edwards, WestEd/Carnegie Math Pathways.
While it is the case that lecture is still the most common form of undergraduate mathematics instruction, research has shown that many of these lectures include a wide range of instructional practices. For instance, one study found that even those instructors who report teaching their course in a lecture format also report incorporating more active techniques, including: “students explaining their thinking, small group work, whole-class discussions, student presentations, and individual work” (Johnson, Keller & Fukawa-Connelly, 2018, pp. 269-70). Additionally, case studies on the pedagogical reasoning of mathematicians often reveal rich belief systems and careful considerations regarding the pedagogical actions that they chose (e.g. Fukawa-Connelly & Newton, 2014; Lai & Weber, 2014; Lew et al., 2016; Nardi, 2007; Weber, 2004). Thus, when discussing where the field is and where it is going, we must adopt a broad view of instruction, one in which lecturing and active learning are not viewed as mutually exclusive by the practitioners. We must also approach this work in a way that positions mathematics instructors not as researchers who are forced to teach, but as reflective and thoughtful educators.

The following two sections address the over-arching questions provided for the disciplinary review of undergraduate mathematics instruction: (1) To what extent, and in what ways, is research-based reform in undergraduate STEM instruction occurring in mathematics? How do you know? and (2) What factors and processes have influenced the depth, breadth and impact of these changes?

**RESEARCH-BASED REFORMS**

The discipline of mathematics has a long history of documentation and data collection on undergraduate instruction. For instance, “every five years since 1965, the Conference Board of the Mathematical Sciences (CBMS) has sponsored a national survey of undergraduate mathematical and statistical sciences in the nation’s two- and four-year colleges and universities” (Blair, Kirkman & Maxwell, 2013, p. xix). Additionally, mathematics has a strong educational research sub-discipline, with many research programs investigating instruction at the undergraduate level. As a result, several national studies have been drawn upon to present an overview of mathematics instruction. When possible, data have been presented that captures differing units of analysis and from multiple institutional contexts.

**CLASSROOM PRACTICES**

Surveying only mathematics and statistics departments, the 2015 CBMS survey results (Blair, Kirkman & Maxwell, 2018) provide some idea about the diffusion of these pedagogical
approaches across colleges and universities (as compared to the concentration within any one institution). New to the 2015 survey, the CBMS survey asked department chairs about the use of various forms of active-learning techniques. Importantly, the department chairs were asked if these approaches were used by a faculty member (i.e., at least one member), as compared to how many faculty members (see Figure 6.1).

In 58% of mathematics departments, at least one faculty member was using inquiry-based learning. This percentage is not consistent across institutional types, however; 71% of master’s-granting departments report inquiry-based learning, with 56% and 57% of PhD- and bachelor’s-granting departments confirming inquiry-based strategies, respectively. Similarly, 58% of the departments reported at least one faculty member using a “flipped classroom” model. However, this practice is significantly lower at master’s-granting departments (52%) compared to PhD (61%) and bachelor’s (59%)-granting departments. More broadly, in 66% of mathematics departments at least one faculty member was using activity-based learning (65% of bachelor’s-granting, 71% of master’s-granting, and 64% of PhD-granting departments).

In statistics, 54% of departments have at least one faculty member doing inquiry-based learning, 39% using a flipped classroom model, and 77% doing activity-based learning. Interestingly, 100% of Master’s-granting statistics departments report that at least one faculty member is using activity-based learning (Blair, Kirkman & Maxwell, 2018). Notice that the converse of these statements implies that in about one-third of math departments no one is implementing activity-based learning techniques.

To better understand the teaching practices of individuals, as opposed to those practices present in departments, we can look to a recent publication by Stains et al. (2018). Based on classroom observations, the researchers were able to classify instructor behaviors as didactic (80% or more of class time spent lecturing), interactive lecture (lecture supplemented with student-centered strategies), or student-centered. Observations of about 200 undergraduate mathematics courses revealed nearly equal amounts of didactic and student-centered

Figure 6.1. Use of Pedagogical Strategies in Math Departments

<table>
<thead>
<tr>
<th>Activity</th>
<th>All Math Depts</th>
<th>PhD Math</th>
<th>MA Math</th>
<th>BA Math</th>
<th>All Stat Depts</th>
<th>PhD Stat</th>
<th>MA Stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inquiry based class</td>
<td>58</td>
<td>56</td>
<td>71</td>
<td>57</td>
<td>54</td>
<td>56</td>
<td>45</td>
</tr>
<tr>
<td>Flipped classroom</td>
<td>58</td>
<td>61</td>
<td>52</td>
<td>59</td>
<td>39</td>
<td>35</td>
<td>55</td>
</tr>
<tr>
<td>Class conducted largely online</td>
<td>38</td>
<td>49</td>
<td>53</td>
<td>33</td>
<td>48</td>
<td>49</td>
<td>45</td>
</tr>
<tr>
<td>Activity based learning</td>
<td>66</td>
<td>64</td>
<td>71</td>
<td>65</td>
<td>77</td>
<td>70</td>
<td>100</td>
</tr>
<tr>
<td>Technology used to develop conceptual understanding</td>
<td>86</td>
<td>82</td>
<td>91</td>
<td>86</td>
<td>84</td>
<td>84</td>
<td>82</td>
</tr>
</tbody>
</table>

courses, approximately 40% of each, with the remaining 20% being classified as interactive lecture.

Unlike most large-scale studies describing undergraduate instruction, it is important to note that the Stains et al. (2018) report was based on observation data, as opposed to surveys with self-reported data. Additionally, the instructional classifications were clearly defined, as opposed to other studies in which terms such as “inquiry” and “class discussions” are more subjective. With those caveats in mind, we now turn our attention to the most recent Higher Education Research Institute’s (HERI) surveys, which were self-reported and open to interpretation. The most recent HERI report analyzed data collected in 2014. For mathematics, these data came from 668 faculty at four-year institutions, representing a national sample of full-time mathematics faculty who teach undergraduate students (Eagan, 2016).

The HERI study found that about two-thirds of mathematics teachers use “extensive” lecturing in all or most of their courses (with 34.7% of mathematics faculty reporting extensive lecturing in all of their courses). When considering more reform-oriented teaching practices, the 2014 HERI survey found that roughly two in five faculty in mathematics (39.1%) utilized class discussions in all of their courses, and almost 20% more used class discussions in most of their courses. In fact, only about 15% of mathematics faculty stated that they used whole-class discussions in none of their courses. Further, in regards to student inquiry, one in five (19.8%) reported using student inquiry to drive learning in all of their courses, a little more than 20% reported using it in most of their courses, and fewer than 25% reported no use of inquiry (Eagan, 2016).

In terms of assessment, the 2014 HERI survey found that more than 70% of mathematics instructors did not use electronic quizzes with immediate feedback (e.g., clickers) in any of their courses. Similarly, the use of curve grading was relatively infrequent, with fewer than 10% using a grading curve in all of their courses and about 55% using it in none of their courses (Eagan, 2016).

Interestingly, between the 2010 and 2014 HERI surveys, there was a noticeable shift in the teaching of undergraduate mathematics (Figure 6.2). Faculty reports of extensive lecturing in all or most of their courses fell by about 10 percentage points between 2010 and 2014. Reports of faculty using class discussions in all or most of their courses increased by more than 10 percentage points (after hitting a low point in 2010), and faculty reports of using inquiry to drive student learning in all or most of their courses rose by more than 20 percentage points (Eagan, 2016). Possible reasons explaining this shift will be discussed in subsequent sections.

While the CBMS and HERI surveys provide large-scale data on mathematics instruction as a whole, they do not provide information on particular courses, or how those courses may differ from each other. Thus, we will now shift focus to more course-specific information to provide a more detailed view of undergraduate mathematics instruction.

First-Year Mathematics

First-year mathematics courses occupy a somewhat peculiar place in mathematics education. These three courses overwhelmingly have the highest student enrollments – with 1,445,000 students taking an “introductory level (including precalculus)” course and another 959,000 taking a “calculus level” course in 2015 – yet only 157,000 students in all other courses above the calculus level (CBMS, 2015). Thus, when we talk about the undergraduate
mathematics courses experienced by students, we are almost exclusively talking about these early courses.

In the spring of 2010, the Characteristics of Successful Programs in College Calculus (CSPCC) project\(^1\) team gathered data from a stratified random sample of US colleges and universities, including two-year colleges (Bressoud, Mesa & Rasmussen, 2015). In total, 212 colleges and universities responded, with more than 700 instructors and more than 14,000 students providing information about their experiences in Calculus I. On the end-of-term survey, 493 instructors provided information about the amount of time spent lecturing with 274 (55.6%) reporting “very often,” and another 188 reporting “often” (38.1%). Thus, we see more than 93% of instructors reporting significant use of lecture in Calculus I.

When we look at reports of more active forms of pedagogy within the CSPCC study, we find 26.2% instructors asking students to work together, 23.8% holding whole-class discussions, and .2% having students give presentations “very often” or “often.” However, when asked to characterize their teaching on a four-point scale from “very innovative” to “very traditional,” 61.0% of instructors rated themselves on the innovative side of the continuum. This disconnect, between their reports of active pedagogy and the characterization of their instruction as “innovative,” suggests the need for further research into the aspects and components of teaching that mathematics instructors see as innovative.

In the spring and summer of 2015, another large-scale national survey was conducted to understand first-year mathematics courses, this time for pre-calculus, Calculus I, and Calculus II. The Progress through Calculus (PtC) survey was sent to department chairs at all master’s- and PhD-granting mathematics departments, with 59% (89/152) of the MA/MS-granting departments and 75% (134/178) of the PhD-granting departments responding to at least some portion of the survey (see Apkarian & Kirin [2017] for the technical report on the PtC survey).

\(^1\) For program description, see https://www.maa.org/cspcc

---

**Figure 6.2.** Student Inquiry Trends: Percentage of classes reporting use of student inquiry, by year

In pre-calculus, 177 departments provided information for 262 courses that functioned as direct preparation for single-variable calculus (i.e., some departments offer multiple pathways into Calculus I). Figure 6.3 summarizes the instructional practices reported in pre-calculus. In these courses, lecturing while answering students’ questions was the most commonly reported instructional format, with department chairs reporting this was the primary mode of instruction in 58.6% of pre-calculus courses.2 In terms of active learning, 18.4% of courses were categorized as lecture with some active learning and another 3.9% as minimal lecture with mainly active learning techniques. For those respondents who reported at least some active learning in pre-calculus, information on 53 courses was provided to describe which active learning techniques are used. POGIL (Process Oriented Guided Inquiry Learning) is used in three pre-calculus courses, Inquiry-Based Learning (IBL) in nine, clicker surveys in 11, group work in 43, and flipped classes in 14.

In Calculus I (Figure 6.4), we see an increase in the use of lecture, with 65.3% of courses primarily using “lecture and answering student questions,” 17% using lecture with some active learning, and 2.8% using minimal lecture with mainly active learning. In the 59 courses that used at least some active learning, three used POGIL, 10 used IBL, 13 used clicker surveys, 50 used group work, and 14 used flipped classes.

As we move into Calculus II (Figure 6.5), we again see an increase in the use of lecture. “Lecture while answering student questions” now describes 73.5% of the courses, with 12.8% being described as lecture with some active learning, and just 1% described as

---

2 Notice that, for the Progress through Calculus survey, the unit of analysis here is courses (e.g., “Pre-Calculus with Trigonometry”); each of these courses typically has multiple sections taught by multiple instructors. Department chairs were asked to comment on the course, not individual sections. They were provided a “there is too much variation” response option, selected for 19/256 courses.
minimal lecture with mainly active learning. In the 41 courses that used at least some active learning, one used POGIL, seven used IBL, 10 used clicker surveys, 33 used group work, and six used flipped classes.

As previously noted, the vast majority of students experience mathematics only through these introductory courses. Additionally, the vast majority of students in first-year mathematics courses are not math majors, with only 1% in the CSPCC survey reporting that they were intending to major in mathematics. Thus, in the introductory courses, instructors are teaching mathematics to almost 2.4 million non-math majors. It is conceivable that in these (often heavily coordinated) courses, which serve so many students and try to meet the needs of numerous client disciplines, instructors are faced with constraints on their instructional practice. Thus, to provide a fuller picture of mathematics instruction, including less constrained settings, we next consider data collected in upper-division mathematics courses.

**Upper-Division Mathematics**

As argued by Johnson, Keller, and Fukawa-Connelly (2018), in upper-division mathematics courses many of the routinely cited obstacles to non-lecture pedagogy are alleviated. For instance, these courses are normally taught to mathematics majors in small classes (typically less than 40 students). While the students have taken numerous pre-requisites, there are only a handful of courses that would serve as a follow-up course. Thus, a certain amount of mathematical sophistication and success can be assumed and there are fewer concerns about preparing students for “the next course.” It is common for at most a handful of sections to run per year, so coordination is minimal, providing autonomy for instructors to have control over pedagogical decisions. Finally, these courses, more often than not, are taught by tenured or tenure-track professors who have job security and autonomy – as opposed to lower-division courses that typically are taught by non-tenure-track instructors and graduate teaching assistants. Thus, there is reason to believe teaching practices in upper-division courses may vary from those in lower-division courses.

In 2015/2016 a survey was sent to abstract algebra instructors, focusing on teaching practice, individual characteristics (e.g., demographic information, beliefs about teaching, beliefs about students, and interest in professional activities), and situational characteristics (e.g., course/curriculum

---

For example, the CSPCC study found 30% of Calculus I students are majoring in engineering, 30% in biology, 8% in business, and 6% in physical science.

---

3 In 2010, the last year the data was available from the CBMS survey, 78% of upper-division mathematics course were taught by tenured or tenure-track professors, as opposed to 32% at the “introductory level” and 59% at the “calculus level.”
information, perceptions of departmental support), and knowledge of and openness to non-lecture practices. In total, there were 219 respondents: 96 from bachelor’s-granting institutions, 44 from master’s-granting institutions, and 79 from PhD-granting institutions. Just focusing on the reported percentage of class time lecturing (shown in Table 6.1), Johnson, Keller, Peterson and Fukawa-Connelly (2018) found that 26% of instructors reported lecturing more than 75% of the time, 57% reported lecturing between 25-75% of the time, and 17% reported lecturing less than 25% of the time.

In this study, Johnson, Keller, Peterson and Fukawa-Connelly (2018) were able to provide detailed information that helps to flesh out the pedagogical practices of these three groupings of instructors. Those who lecture less than 25% of the time reported using the rest of their class time (split fairly evenly) on showing students how to write proofs, having students work in small groups, having students give presentations, having students work individually, lecturing, holding whole class discussions, and having students explain their thinking. Students in these classes are frequently asked to make presentations to the class and develop their own conjectures and proofs, and are sometimes asked to develop their own definitions. Those who lecture between 25-75% of the time, said they spend a significant amount of class time showing students how to write specific proofs, pausing to ask students questions, and using diagrams, visual representations, and informal explanations to help students with formal ideas. There is some class time devoted to students working alone and in small groups, giving presentations, and explaining their thinking, and students are pretty frequently asked to develop their own proofs. Finally, those who lecture for more than 75% of the time report they are also showing students how to write specific proofs and pausing to ask students questions. These lectures often include diagrams to illustrate ideas and informal explanations of formal statements. Students in these courses are sometimes asked to develop their own conjectures or proofs.

**Conclusions**

There are some inconsistencies when considering these national survey studies as a whole, inconsistencies that are exacerbated due to the fact that surveys each used similar but not identical survey items and had differing units of analysis (e.g., instructors for HERI, departments for CBMS, and courses for PtC). Here, we try to reconcile some of these discrepancies.

First, the HERI study found that nearly 45% of mathematics faculty reported using student inquiry to drive learning in all or

<table>
<thead>
<tr>
<th>Institutional type</th>
<th>Number of instructors reporting extent of lecture</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Terminal math degree offered</td>
<td>Lecture less than 25%</td>
<td>Lecture between 25-75%</td>
</tr>
<tr>
<td>Bachelor’s Degree</td>
<td>19</td>
<td>65</td>
</tr>
<tr>
<td>Master’s Degree</td>
<td>12</td>
<td>21</td>
</tr>
<tr>
<td>PhD</td>
<td>7</td>
<td>39</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>38</strong></td>
<td><strong>125</strong></td>
</tr>
</tbody>
</table>

(Data from Johnson, Keller, Peterson, and Fukawa-Connelly, 2018)
most of their courses. However, only 17% of Abstract Algebra instructors reported minimal lecture (i.e., less than 25% of class time). Further, mainly active learning was only reported in 3.9% of pre-calculus courses, 2.8% of Calculus I courses, and only 1% of Calculus II. Given the course-specific data, it appears that mathematics faculty do not view “student inquiry that drives learning” and “minimal lecture” as equivalent instructional approaches – nor are they viewing student inquiry as antithetical to predominantly lecture-based instruction. Instead, these studies suggest that, when mathematics instructors are bringing “student inquiry” and “some active learning techniques” into their instruction, they are incorporating these approaches into their lectures, rather than using them instead of lecturing.

Second, the HERI study reported that about two thirds of mathematics teachers use “extensive” lecturing in all or most of their courses. While this is a significant proportion, it is lower than that being reported at the course level. In the lower-division courses, 77% of pre-calculus courses, 82.3% of Calculus I courses, and 86.3% of Calculus II courses are taught as either “lecturing and answering student questions” or “lecture with some active learning techniques.” The mismatch between rates of lecture in these courses and rates of lecturing by mathematics faculty in the HERI study could stem from any number of factors – most obviously, the HERI survey asked about faculty and the PtC survey asked about courses and it is unclear if “lecture with some active learning techniques” counts as “extensive lecturing.” However, the large population of instructors and students in these courses warrants drawing attention to this mismatch. At four-year colleges and universities, 92.8% of students enrolled

in mathematics courses are enrolled in “calculus level” courses or lower. Thus, if we are talking about what students experience in undergraduate mathematics courses, these courses must carry much more weight in our discussions. Additionally, these courses are the least likely to be taught by tenured or tenure-track professors (or even full-time faculty). We cannot ignore those who teach lower-division courses (e.g., part-time adjuncts and graduate teaching assistants) when discussing instructional change in undergraduate mathematics.

**REFORMS IN INSTRUCTOR INTEREST, BELIEFS, AND ATTITUDES**

The mathematics professional societies have recently taken a strong stance on the state of undergraduate mathematics education, stating unequivocally that “the status quo is unacceptable.” This statement came out of the *Common Vision* project (Saxe & Braddy, 2015), which included representatives from the American Mathematical Association of Two-Year Colleges (AMATYC), the American Mathematical Society (AMS), the American Statistical Association (ASA), the Mathematical Association of America (MAA), and the Society for Industrial and Applied Mathematics (SIAM). The *Common Vision* project synthesized seven curricular guides, which had been published by these five societies, in order to identify and report on common themes. In regards to pedagogy, *Common Vision* states that:

*Across the guides we see a general call to move away from the use of traditional lecture as the sole instructional delivery method in undergraduate mathematics courses.... Even within the traditional lecture setting, we should seek to more actively engage students than we have in the past* (Saxe & Braddy, 2015, p. 19).

It appears that mathematics departments are also seeing the value in active-learning techniques. In the PtC survey, department chairs and/or course coordinators were asked how important active learning is for having a successful Pre-calculus – Calculus II sequence. When aggregated, the data from both PhD-granting and Master’s-granting institutions showed that 44.3% of institutions felt that active learning was “very” important.

Figure 6.7. Success of features in Pre-calculus through Calculus II programs: Proportion of departments reporting the success of each feature, by institutional type

<table>
<thead>
<tr>
<th>Features</th>
<th>All</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Very</td>
<td>Some</td>
<td>Not</td>
<td>N</td>
<td>Very</td>
<td>Some</td>
<td>Not</td>
<td>N</td>
<td>Very</td>
<td>Some</td>
<td>Not</td>
<td></td>
</tr>
<tr>
<td>Challenging courses</td>
<td>214</td>
<td>0.425</td>
<td>0.514</td>
<td>0.061</td>
<td>130</td>
<td>0.408</td>
<td>0.508</td>
<td>0.085</td>
<td>84</td>
<td>0.452</td>
<td>0.524</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td>Uniform components</td>
<td>210</td>
<td>0.624</td>
<td>0.352</td>
<td>0.024</td>
<td>127</td>
<td>0.701</td>
<td>0.283</td>
<td>0.016</td>
<td>83</td>
<td>0.506</td>
<td>0.458</td>
<td>0.036</td>
<td></td>
</tr>
<tr>
<td>Instructor meetings</td>
<td>195</td>
<td>0.215</td>
<td>0.503</td>
<td>0.282</td>
<td>119</td>
<td>0.277</td>
<td>0.479</td>
<td>0.244</td>
<td>76</td>
<td>0.118</td>
<td>0.539</td>
<td>0.342</td>
<td></td>
</tr>
<tr>
<td>Monitoring local data</td>
<td>212</td>
<td>0.179</td>
<td>0.599</td>
<td>0.222</td>
<td>128</td>
<td>0.188</td>
<td>0.602</td>
<td>0.211</td>
<td>84</td>
<td>0.167</td>
<td>0.595</td>
<td>0.238</td>
<td></td>
</tr>
<tr>
<td>Student placement</td>
<td>215</td>
<td>0.386</td>
<td>0.586</td>
<td>0.028</td>
<td>129</td>
<td>0.280</td>
<td>0.605</td>
<td>0.016</td>
<td>86</td>
<td>0.395</td>
<td>0.558</td>
<td>0.047</td>
<td></td>
</tr>
<tr>
<td>GTA preparation</td>
<td>185</td>
<td>0.341</td>
<td>0.503</td>
<td>0.157</td>
<td>127</td>
<td>0.362</td>
<td>0.528</td>
<td>0.110</td>
<td>58</td>
<td>0.293</td>
<td>0.448</td>
<td>0.259</td>
<td></td>
</tr>
<tr>
<td>Student support programs</td>
<td>216</td>
<td>0.421</td>
<td>0.556</td>
<td>0.023</td>
<td>130</td>
<td>0.400</td>
<td>0.577</td>
<td>0.023</td>
<td>86</td>
<td>0.453</td>
<td>0.523</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>Active learning</td>
<td>199</td>
<td>0.151</td>
<td>0.668</td>
<td>0.181</td>
<td>117</td>
<td>0.128</td>
<td>0.658</td>
<td>0.214</td>
<td>82</td>
<td>0.183</td>
<td>0.683</td>
<td>0.134</td>
<td></td>
</tr>
</tbody>
</table>

– with only 9.1% indicating that active learning was “not” important. Figure 6.6 shows how the perceived importance of active learning compares to other features that departments reported.

Respondents were then asked how successful they were at implementing each of these factors. While 44.3% of department chairs said active learning was very important, only 15.1% thought they were very successful in implementing it– with another 66.8% reporting that they were somewhat successful (see Figure 6.7).

Zooming in to consider individual instructors, there are two national surveys that provide data on instructors’ beliefs on instruction: CSPCC, administered to Calculus I instructors, and Johnson, Keller, Peterson, and Fukawa-Connelly’s (2018) survey of abstract algebra instructors. In the CSPCC study, at the end of the semester instructors were asked a few items regarding their beliefs related to teaching. Among respondents, 65% agreed, to various degrees, that calculus students learn best from lecture and 85% agreed, to various degrees, that research literature on how students think about calculus ideas would be useful for teaching (Table 6.2).

These instructors were also asked survey questions about their interest around professional development, with 37% of instructors reporting they were very interested in raising their awareness of how student learn calculus ideas, and 65% reporting a very strong interest in improving their teaching (Table 6.3).

Johnson, Keller, Peterson and Fukawa-Connelly’s (2018) survey, administered to instructors of abstract algebra (an upper-division mathematics course) asked several Likert-scale items on instructional beliefs (on a four-point scale ranging from “strongly disagree” to “strongly agree.” (Table 6.4).

Two of the survey items appear to be very consistent. About 60% of instructors agree with the statements “I think lecture is the best way to teach” and “I think students learn better if I first explain the material to them and then they work to make sense of the ideas for themselves.” However, an overwhelming majority of instructors agreed with the statements “I think students learn better when they do mathematics work (in addition to taking notes and attending to the lecture) in class” and “I think students learn better when they struggle with the ideas prior to me explaining the material to them” (87% 

<table>
<thead>
<tr>
<th>Survey Item</th>
<th>Strongly disagree</th>
<th>Disagree</th>
<th>Somewhat Disagree</th>
<th>Somewhat Agree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculus students learn best from lectures, provided they are clear and well-organized (n = 470)</td>
<td>3.2%</td>
<td>11.7%</td>
<td>20.2%</td>
<td>37.4%</td>
<td>20.0%</td>
<td>7.4%</td>
</tr>
<tr>
<td>Familiarity with the research literature on how students think about ideas in calculus would be useful for teaching. (n = 472)</td>
<td>1.9%</td>
<td>4.2%</td>
<td>8.5%</td>
<td>28.4%</td>
<td>39.0%</td>
<td>18.0%</td>
</tr>
</tbody>
</table>

(Data for this table can be requested at: https://www.maa.org/programs/faculty-and-departments/curriculum-development-resources/national-studies-college-calculus/data-for-researchers.)
agreement and 77% agreement, respectively). These responses are somewhat contradictory but promising for active learning.

Taken together, this collection of beliefs may be consistent with a mode of instruction in which students are expected to read the textbook before attending class (students struggle with ideas prior to explanation), the instructor then lectures (lecture is best) with opportunities for individual seat work on practice problems (students do mathematical work in class), followed by a homework set (I first explain the material and then they work to make sense of the ideas themselves). Thus, while some of these survey items may appear to be promising when considered on their own, they may actually reflect a mode of instruction that is more lecture-based than active-learning.

**Conclusions**

Taking these studies and reports together, it appears that there has been (at least in principle) a shift in values in the professional community. That stance appears to be reflected at the department level, with department chairs and course coordinators reporting some positive beliefs about the importance of active learning. However, individual instructors still report strong beliefs in the strength and utility of lecture. Perhaps on a promising note, instructors also report strong interest in professional development and views on student learning that may be leveraged to develop more active pedagogies.

### Table 6.3. CSPCC Interest in Improving Instruction: Percentage of instructors reporting interest.

<table>
<thead>
<tr>
<th>Survey Item</th>
<th>Very strong</th>
<th>Moderately strong</th>
<th>Mildly strong</th>
<th>Not at all</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participating in activities that raise your awareness of how students learn key ideas in calculus?</td>
<td>36.9%</td>
<td>39.0%</td>
<td>20.0%</td>
<td>4.1%</td>
</tr>
<tr>
<td>Improving your own teaching?</td>
<td>64.8%</td>
<td>27.4%</td>
<td>6.7%</td>
<td>1.1%</td>
</tr>
</tbody>
</table>

(Data for this table can be requested at: [https://www.maa.org/programs/faculty-and-departments/curriculum-development-resources/national-studies-college-calculus/data-for-researchers.](https://www.maa.org/programs/faculty-and-departments/curriculum-development-resources/national-studies-college-calculus/data-for-researchers.))

### Table 6.4. Belief Items on Abstract Algebra Survey: Percentage of instructors reporting belief

<table>
<thead>
<tr>
<th>Survey Item</th>
<th>Disagree</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I think lecture is the best way to teach. (n = 217)</td>
<td>41%</td>
<td>59%</td>
</tr>
<tr>
<td>I think lecture is the only way to teach that allows me to cover the necessary content. (n = 214)</td>
<td>47%</td>
<td>53%</td>
</tr>
<tr>
<td>I think students learn better when they do mathematical work (in addition to taking notes and attending to the lecture) in class. (n = 216)</td>
<td>13%</td>
<td>87%</td>
</tr>
<tr>
<td>I think students learn better when they struggle with the ideas prior to me explaining the material to them. (n = 215)</td>
<td>23%</td>
<td>77%</td>
</tr>
<tr>
<td>I think students learn better if I first explain the material to them and then they work to make sense of the ideas for themselves. (n = 217)</td>
<td>40%</td>
<td>40%</td>
</tr>
</tbody>
</table>

Generally speaking, there is no standard teaching preparation offered in mathematics. However, many mathematics instructors have likely participated in some sort of preparation to become a Graduate Teaching Assistant (GTA). As reported in the Progress through Calculus (PtC) survey, researchers found 52% of department chairs and/or course coordinators reported GTA preparation to be “very” important to a successful Pre-calculus through Calculus II program, with another 33% reporting it was “somewhat” important (Apkarian & Kirin, 2017).

Of the participants in the PtC study, 83% of PhD-granting, and 47% of master’s-granting mathematics departments have a required GTA teaching preparation program. (At 48% of institutions there was also university-wide GTA preparation, i.e., not math-specific and not directed by the mathematics department.) However, there is wide variability in these programs, both in terms of length and in terms of in substance. As reported in Figure 6.8, 57% of departments reported that their GTA teaching program was a year-long course or seminar, whereas 28% reported they had a short workshop or orientation, and 15% said they only have one day of preparation for new GTAs (Apkarian & Kirin, 2017).

The most commonly reported activities in these GTA preparation programs were developing lesson plans (41%); learning about assessment methods (40%); and watching someone else teach (or reading vignettes) (34%). Additionally, feedback in the form of mentoring and observations appear to be a significant component to many GTA preparation programs. A common feature for these programs was formal feedback for the GTA’s teaching. As evident in Figure 6.9 (with a sum of 233% in the “All” column), many programs provide multiple forms of feedback (Apkarian & Kirin, 2017). Overall, departments appear to be fairly satisfied with their GTA preparation programs. When asked, “How well does your teaching preparation program prepare new GTAs for their roles in the precalculus/calculus sequence?” 21% responded “very well,” 39% responded “well,” and 39% responded “adequately.” Only one institution out of 140 stated that their program did “poorly” or “very poorly” at preparing GTAs (Apkarian & Kirin, 2017).

There are a number of other research programs investigating GTA preparation, including the work of Natasha Speer. There is room for further discussion around what is known about the structure, scope, and effects of GTA preparation programs. Of particular interest would be any research that may speak to the
viability of such programs as an effective route to instructor uptake of active learning. After the completion of a GTA preparation program, opportunities for professional development dwindle. Notable exceptions include the MAA’s Project NExT (New Experiences in Teaching) and the workshops offered through the Academy of Inquiry-Based Learning. Project NExT is intended for newly graduated or recently appointed PhDs and is focused on “improving the teaching and learning of mathematics, engaging in research and scholarship, finding exciting and interesting service opportunities, and participating in professional activities” (Project NExT, n.d.). Since 1994 there have been 1700 fellows, with about 80 new fellows in each yearly cohort (LaRose, 2018). These fellows attend a three-day workshop and a follow-up meeting one year later. Discussions at these workshops and meetings include topics around pedagogy, such as: innovative teaching approaches, strategies for engaging students, and supporting underserved and under-represented groups of students. However, these workshops serve as more of a survey of issues specifically relevant to new faculty, as opposed to a deep and focused professional development program.

The Academy of Inquiry Based Learning does offer such a program. Inquiry-based learning (IBL) is often characterized as an instructional paradigm that prioritizes students’ active engagement in mathematical activities during class. This could include several different instructional techniques, such as student presentations and small group work but, more pivotal, “students actively participate in contributing their mathematical ideas to solve problems,

### Figure 6.9. GTA Preparation Program Feedback.

Which of the following activities, related to providing feedback on GTA’s teaching, does your program formally include? Mark all that apply.

<table>
<thead>
<tr>
<th>Activity</th>
<th>All (156)</th>
<th>PhD (112)</th>
<th>MA (44)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTAs practice teaching and receiving feedback on their teaching</td>
<td>105 (0.673)</td>
<td>83 (0.741)</td>
<td>22 (0.500)</td>
</tr>
<tr>
<td>GTAs are observed by an experienced instructor while teaching in the classroom and receive feedback on their teaching</td>
<td>117 (0.750)</td>
<td>85 (0.759)</td>
<td>32 (0.727)</td>
</tr>
<tr>
<td>New GTAs are observed by an experienced instructor while teaching in the classroom and receive feedback on their teaching</td>
<td>41 (0.263)</td>
<td>37 (0.330)</td>
<td>4 (0.091)</td>
</tr>
<tr>
<td>New GTAs teaching in the classroom are videotaped for review and discussion with a mentor or experienced instructor</td>
<td>22 (0.141)</td>
<td>22 (0.196)</td>
<td>0 (0.000)</td>
</tr>
<tr>
<td>GTAs are paired with a mentor to discuss teaching</td>
<td>56 (0.359)</td>
<td>39 (0.348)</td>
<td>17 (0.386)</td>
</tr>
<tr>
<td>Other</td>
<td>11 (0.071)</td>
<td>8 (0.071)</td>
<td>3 (0.068)</td>
</tr>
<tr>
<td>No response</td>
<td>12 (0.077)</td>
<td>6 (0.054)</td>
<td>6 (0.136)</td>
</tr>
</tbody>
</table>

rather than applying teacher-demonstrated techniques to similar exercises” (Yoshinobu & Jones, 2012, p. 307). The Academy of IBL offers intensive, four-day, summer workshops, with morning sessions devoted to lesson studies and discussions of mathematics education research and afternoon sessions devoted to providing participants supports for developing their own curricula materials. Since 2006, 368 mathematics instructors have participated in one of the Academy of IBL summer workshops (Stan Yoshinobu, personal communication, March 14, 2018).

The research that has been done on the effects of such professional programs appears to show that they hold promise for influencing instruction. For instance, Hayward, Kogan, and Laursen (2016) conducted a study of 139 instructors who attended an IBL workshop. They found that, in the year following the workshop, 58% of the instructors reported implementing IBL instructional strategies. Additionally, while the number of instructors attending these workshops is relatively small in comparison to the large number of mathematics instructors, small numbers may not mean influential numbers. As we will see in later sections, advocacy of individual instructors within a department is a highly reported lever for change.

DIVERSITY AND INCLUSION

Data suggest that there is some attention to diversity and inclusion by mathematics faculty. The 2014 HERI survey included an item asking if faculty used techniques to create inclusive classrooms for diverse students, and 21.2% of mathematics faculty reported using inclusive classroom techniques in all of their courses. However, 36.3% reported not using inclusive techniques in any of their courses (Eagan, 2016). Given the nature of the study (i.e., closed-form survey), it is impossible to determine what these faculty members believed to be “techniques to create inclusive classrooms for diverse students.”

In June 2016, the MAA held a conference Precalculus to Calculus: Insights and Innovations. At this conference 104 participants (representing 56 U.S. colleges and universities), came together to discuss issues facing Pre-calculus – Calculus II instruction. In reflecting on the themes emerging at this conference, Apkarian, Kirin, Gehrtz, and Vroom (2017) concluded:

Participants recognized the moral and economic imperatives for addressing the particular issues faced by women, students of color, first-generation college students, and other underrepresented groups in STEM.... It became apparent, however, that attendees underappreciated how much action undertaken within the mathematics department can make a difference to these students. Many institutions have support systems in place for these students, but few mathematics departments consciously integrate this support into mathematics programs or educate their faculty about topics such as stereotype threat or inclusive teaching (p. 36).

Overall, much more work needs to be done in this area, both in regard to inclusive teaching practices (e.g., what is effective, how knowledgeable are faculty regarding inclusive mathematics instruction, and how prevalent are these practices) and in terms of inclusive mathematics departments and programs.

FACTORS AND PROCESSES

As previously noted, the 2014 HERI report showed an uptick in active learning from 2010 to 2014. This report is corroborated
by the 2015 CBMS survey, in which 60% of mathematics departments and 80% of statistics departments reported major pedagogical changes in the last 10 years (Blair, Kirkman & Maxwell, 2018; see Figure 6.10). Here we look at the reported influences on, and from, the mathematics professional societies, departments, and faculty members.

In the statement prepared by the Common Vision committee, which included representatives from all the five of the national mathematical professional societies, two reports were identified as particularly influential: The President’s Council of Advisors on Science and Technology’s Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics report (Olson & Riordan, 2012) and The National Research Council’s The Mathematical Sciences in 2025 report (NRC, 2013). Common Vision highlighted two aspects of the PCAST report, “dissatisfaction in how undergraduate mathematics is taught to students outside the mathematics major” and “outdated course materials and teaching techniques have not provided students with the quantitative skills demanded for employment and good citizenship” (p. v). In terms of the NRC report, Common Vision singled out the call “for mathematics teaching that better aligns with other disciplines” (p. v).


**Figure 6.10. Departments Reporting Major Changes**

<table>
<thead>
<tr>
<th>Activity</th>
<th>All Math Depts</th>
<th>PhD Math</th>
<th>MA Math</th>
<th>BA Math</th>
<th>All Stat Depts</th>
<th>PhD Stat</th>
<th>MA Stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Department experienced major changes over the last 10 years</td>
<td>60</td>
<td>62</td>
<td>65</td>
<td>58</td>
<td>80</td>
<td>78</td>
<td>85</td>
</tr>
<tr>
<td>Of those experiencing change, the percent attributing the change to:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educational research</td>
<td>61</td>
<td>67</td>
<td>77</td>
<td>56</td>
<td>49</td>
<td>53</td>
<td>36</td>
</tr>
<tr>
<td>Advocacy of some faculty member in the department</td>
<td>91</td>
<td>99</td>
<td>90</td>
<td>90</td>
<td>88</td>
<td>88</td>
<td>91</td>
</tr>
<tr>
<td>Advocacy by another department</td>
<td>16</td>
<td>23</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>Advocacy by institution’s administrators</td>
<td>37</td>
<td>47</td>
<td>30</td>
<td>35</td>
<td>47</td>
<td>48</td>
<td>45</td>
</tr>
<tr>
<td>Advocacy by a professional organization</td>
<td>39</td>
<td>31</td>
<td>33</td>
<td>43</td>
<td>38</td>
<td>36</td>
<td>45</td>
</tr>
</tbody>
</table>

(Reprinted with permission.)
This Instructional Practices Guide is designed as a “how to” guide focused on mathematics instruction at the undergraduate level. It is based on the concept that effective teaching is supported by three foundational types of practices: classroom practices, assessment practices, and course design practices all informed by empirical research as well as the literature on technology and equity (p. 5, emphasis in original).

Thus, it appears there is a growing consensus among, and within, the professional mathematics societies that educational reform is needed in undergraduate mathematics. There is a call for more active classroom engagement that attends to the education research literature and acknowledges equity issues.

There is some indication that this advocacy is informing instruction. As part of the 2015 CBMS survey, mathematics and statistics departments (at four-year colleges and universities) were asked if the department had experienced major change in the types of pedagogy used in the department during last 10 years (see Figure 6.10). Roughly 60% of mathematics departments and 80% of statistics departments reported that there had been major changes in regard to teaching. Those who reported significant changes were then provided a list of potential factors and were asked if any were influential to the changes. The overwhelming factor reported, for both mathematics and statistics, was “the advocacy of some member of their faculty” – selected by 91% of mathematics departments and 88% of statistics departments. Educational research was the next most commonly cited factor, (49% and 61%, respectively), followed by administrators’ advocacy (37% and 47%, respectively).

These findings highlight the importance of key individuals as change agents within departments. Given the reported influence of “advocacy of some faculty member in the department,” to understand instructional change we must understand the influences on individual faculty members. Reporting on their survey of upper-division mathematics instructors in graduate degree-granting...

Figure 6.11. Influences on Instruction of Upper-level Mathematics Courses

![Image of a bar chart](Image)
departments, Fukawa-Connelly, Johnson, and Keller (2016) found that, when asked to report if any of the listed sources were influential on their teaching practice, “experiences as a teacher (84 percent) and experiences as a student (64 percent) were far and away the most significant” (p. 280). The relative reports of influence for the other sources listed are provided below in Figure 6.11.

These findings help to contextualize the scope of the impact that professional development opportunities and outreach initiatives, such as the efforts of MAA’s Project NExT and the Academy of IBL, are having on undergraduate mathematics instruction. Overall:

“[l]ittle importance was assigned to the normal means that grant-supported projects use to disseminate new teaching ideas: Project NExT (8 percent), MathFest, MAA mini-courses or other workshops (13 percent), or publications about teaching such as the MAA Notes series or PRIMUS (2 percent) (p. 280).”

Thus, while it does not appear that these initiatives have a very wide reach, individual advocacy is very influential within departments (Blair, Kirkman & Maxwell, 2018). However, with Project NExT and the Academy of IBL directly working with +100 faculty members a year, the influence of these communities on individuals, and these individuals on their departments, warrants further investigation.

An additional source of influence on undergraduate mathematics instruction that warrants further investigation is GTA professional development programs. For instance, the University of Michigan currently has over 60 post-doctoral fellows and over 130 PhD students. Many of these fellows and students participate in extensive professional development activities, including an intensive four-day training seminar that takes place the week before they become course instructors, weekly course meetings throughout the semester, and observations and feedback. Additionally, the courses that these post-doctoral fellows and GTAs teach are highly coordinated and emphasize active learning models for classroom instruction. The long-term impact of this program, and similar ones across the nation, warrants further investigation, especially given the importance instructors place on their own early experiences as a teacher in shaping their pedagogical approaches.

**CONCLUSION**

In regard to the use of reform-oriented instructional practices in undergraduate mathematics, there are some promising findings in the available research. The 2014 HERI survey found that 39.1% of instructors utilized class discussions in all of their courses, and almost 20% more used class discussions in most of their courses (Eagan, 2016). The most recent CBMS study found that in 58% of mathematics departments, at least one faculty member was using inquiry-based learning (Blair, Kirkman & Maxwell, 2018).

---

5 Data accessed from the University of Michigan’s Mathematics Department’s website: https://lsa.umich.edu/math/people/
Stains et al. (2018) found that 40% of the classrooms observed were student-centered and another 20% were interactive lecture. However, drawing implications from these findings is difficult given that other research indicates that lecture is still the dominant form of instruction. For instance, the PtC study found that “mainly active learning” was only reported in 3.9% of pre-calculus courses, 2.8% of Calculus I courses, and only 1% of Calculus II (Apkarian, & Kirin, 2017).

One way to reconcile these conflicting results is to acknowledge that student-centered instruction and lecture, as one aspect of pedagogical practice, are not entirely distinct. Instead, these studies suggest that, when mathematics instructors bringing “student inquiry” and “some active learning techniques” into their instruction, they are incorporating it into their lectures – as opposed to adopting such instruction instead of lecturing. Thus, those working to enact instructional change in undergraduate mathematics must be careful to not impose dichotomous language that does not resonate with the practitioners they are trying to reach.

Additionally, there are some promising findings regarding beliefs and attitudes towards instructional reform. It appears there has been a shift in instructional values in the professional community, as evident by the Common Vision report (Saxe & Braddy, 2015) and the Instructional Practices Guide (Mathematical Association of America, 2018). That stance appears to be reflected at the department level, with the PtC project finding that a significant proportion of department chairs and course coordinators hold positive beliefs about the importance of active learning (Apkarian, & Kirin, 2017). Interestingly, while individual instructors report strong beliefs in the strength and utility of lecture, they also express beliefs
that are consistent with more active instruction (Johnson, Keller, Peterson, & Fukawa-Connelly, 2018). Such beliefs may offer an opening for instructional change, especially if that change is supported through professional development programs, such as those offered through the Academy of Inquiry Based Learning.

While a number of studies have been conducted in undergraduate mathematics education, more research needs to be done. Little is known in regard to instructors’ attention to diversity and inclusion. The effects of GTA preparation programs on instructional practice also warrant further research. Finally, with the CBMS report suggesting the powerful influence of individual instructors (Blair, Kirkman & Maxwell, 2018), more work needs to be done to investigate the interplay between individuals, departmental cultures, and instructional change.

REFERENCES
of Research in Undergraduate Mathematics Education, 2(1), 59-82.
INTRODUCTION

In physics and astronomy, many research-based instructional strategies (RBIS) have been developed and disseminated over the last several decades, with data that documents improved outcomes. Many of these approaches are based on constructivism (Jonassen & Land, 2000) and related theories, where students take an active role in interacting with their environment to build understanding based on previous knowledge and experiences (Bybee, Carlson-Powell & Trowbridge, 2008; Von Glasersfeld, 1989). RBIS often encourage university instructors to move away from passive lecture to active, student-centered approaches, known collectively as “interactive engagement” (Hake, 1998). Often this involves students working in teams on authentic and challenging tasks or problems with instructors providing scaffolding and feedback (Ambrose et al., 2010).

Due to dissemination of RBIS through professional development efforts (e.g. Henderson, 2008), online resources (e.g., McKagan, n.d.), and word of mouth, it appears that awareness of these alternate teaching strategies has spread. For example, Henderson & Dancy (2009) found almost all faculty (87.1%) they surveyed were familiar with one or more RBIS and approximately half were familiar with six or more RBIS. In addition, approximately half of instructors surveyed (48.1%) said that they used at least one RBIS.
There is also evidence that instructors are aware of the shortcomings of the lecture method and believe student-centered practices can be beneficial. For example, Turpen, Dancy and Henderson’s (2016) study revealed that instructors overwhelmingly cite the primary advantage of peer instruction (PI) is that it is not traditional lecture. Additionally, the beliefs of many instructors align with RBIS, such as valuing active learning and opportunities for students to work together (Henderson & Dancy, 2008).

Despite growing instructor awareness and instructional attitudes that align with RBIS, use of reformed teaching has not expanded to the same extent (Walczyk & Ramsey, 2003; Henderson & Dancy, 2009). When asked about the difference between their beliefs about teaching and teaching practice, instructors cite strong situational constraints as interfering with their ability to make desired changes: large class sizes, expectations for broad content coverage, classroom infrastructure, scheduling constraints, poor student preparation or motivation, and student resistance (Henderson & Dancy, 2007). Yet the discrepancy between theoretical best practices and implemented teaching strategies cannot be explained simply as a result of these situational constraints (Henderson & Dancy, 2007; Henderson et al., 2005). Changing instructional behavior also is limited by implicit instructional models that are somewhat at odds with research-based instruction, incomplete knowledge of instructional strategies (Henderson, 2005), and misalignment of dissemination strategies (Henderson & Dancy, 2008; Henderson & Dancy, 2008). Similarly, Dokter (2008) found that astronomy instructors would list external reasons—curriculum and instruction, time, logistics and student response—over personal reasons as barriers to the use of teaching strategies.

Broader institutional change is an even more complex problem, exaggerated by challenges to moving a curriculum beyond
its development site (Sharma et al., 2010; Finkelstein & Pollock, 2005), where developers often had expertise and resources not found at secondary sites. Changing the context of a reform introduces challenges with departmental dynamics and norms, political struggles, institutional structures, the reward system and more (Dancy & Henderson, 2008; Dancy & Henderson, 2005).

This review describes attempts to modernize university physics teaching, starting with course reform efforts, followed by instructional change initiatives, then broader efforts aimed at changing institutions and the discipline as a whole. This paper is not a comprehensive collection of every reformed curriculum and every change strategy. Rather, we chose specific examples across a variety of categories and highlight the successes and challenges of each reform effort. Since colleges and universities are complex systems, a simple “one size fits all” approach will not work; rather, providing examples across diverse contexts illuminates common trends, and some strengths and shortcomings associated with different approaches. We conclude with a summary of which domains are the most well developed with respect to reform and identify areas where progress needs to be made.

**RESEARCH-BASED REFORM**

In this section, we provide examples of several influential RBIS designed for use with (a) lecture, (b) tutorials, (c) laboratories, and, more significant, (d) structural changes. For a more extensive list of RBIS in physics, see Docktor and Mestre (2014) and Henderson and Dancy (2009). The RBIS featured here have spurred secondary implementations and entered the reform literature. Many of these RBIS target introductory quantitative physics, where innovations are often the most widespread and well documented. Many of these RBIS, including peer instruction and tutorials, have been adapted for use in upper-division courses (e.g., Pollock, Chasteen, Dubson & Perkins, 2010) and astronomy (Prather, Rudolph & Brissenden, 2009; Prather, Rudolph, Brissenden & Schlingman, 2009), although fewer publications are available on secondary reform efforts in these areas. Many of the studies come from large research institutions; there is less formal information about what is happening at liberal arts and community colleges. After describing important innovations in each category, case studies from the literature are used to describe how secondary sites use the innovation. Some of the studies survey multiple users (e.g., Henderson & Dancy, 2009) or try to measure the extent of the reform worldwide (e.g., Foote, Neumeyer, Henderson, Dancy & Beichner, 2014).

**METHODS TO AUGMENT LECTURES**

In order for instructors to incorporate more active learning and engaging activities during class, some strategies ask students to prepare for class by reading, watching videos or doing an interactive online activity. Others integrate activities to break up lecture time and to promote interaction with peers as discussed below.

**Pre-lecture Assignments**

One pre-lecture method, called just-in-time teaching (JiTT), requires students to submit answers to open-ended, conceptual questions electronically before class (Novak, Patterson, Gavrin & Christian, 1999; Novak, 2011). In addition to ensuring students come to class prepared, JiTT helps instructors align their instruction with the needs of their class, so they can explain concepts in the words of their students or address difficulties. As a versatile method, JiTT is often paired with
other RBIS, for example with peer instruction (Crouch & Mazur, 2001).

In the reform literature: Since its invention in the 1990s, JiTT has spread beyond the United States and into other courses and disciplines beyond introductory physics (Gavrin, 2006; Novak, 2011). Henderson and Dancy (2009) found that almost half of the instructors they surveyed had heard of JiTT and it was the sixth most used reform, used by about 8% of respondents. The developers of JiTT disseminated the reform largely through traditional means—publications, web sites, and oral presentations (Gavrin, 2006). However, a survey of JiTT users revealed that most people learned about JiTT from colleagues, while conferences and professional development initiatives (Henderson, 2008; Kezar & Elrod, 2012) have also spread awareness (Gavrin, 2006).

Gavrin (2006) describes a very successful dissemination of JiTT to over half of the faculty at a small liberal arts college in South Carolina. According to this report, strong administrative support for JiTT inspired pride and provided funding for professional development. Widespread adoption was also driven by the decision to use JiTT in the first-year seminar taken by all students and taught by all faculty members in rotation. In this way, almost all instructors developed moderate familiarity with the methods, and they often adapted it in teaching other courses.

Implications for change: Wide dissemination of JiTT may largely be due to its flexibility and versatility beyond introductory physics. Dissemination is accelerated by word-of-mouth conversations between colleagues, as well as formally by the developers. This experience agrees with literature on dissemination of innovations, which says that “mass market channels” such as journal publications, websites and conference presentations can raise awareness of an innovation, but interpersonal channels are more influential when it comes to changing behavior (Borrego, Froyd & Hall, 2010; Rogers, 2003; Moore, 2002).

Interactive Lecture Demonstrations
The Interactive Lecture Demonstration (ILD) curriculum describes a structured approach to presenting demonstrations or experiments within a lecture environment. Often these simple experiments—such as running a cart down a track to examine its motion—use computers and sensors to log and display data (Sokoloff & Thornton, 2004). Guided by a worksheet, students make a prediction about what they expect to see, discuss their prediction with peers, observe the event, and compare the observation to their prediction.

In the reform literature: Sharma et al. (2010) describe a decade-long attempt to use ILDs in a multiple-section, large introductory physics class at a research university in Australia. Adopting RBIS in first-year courses is challenging, where multiple instructors in multiple streams may teach hundreds of students using syllabi dictated by departmental committees, so the authors
wanted to see if the benefits of using ILDs justified the upheaval caused by the change. While students’ learning gains on a concept inventory improved over traditional instruction (Sharma et al., 2010), gains improved less than promised by the developers (Sokoloff & Thornton, 2004) and varied from year to year, despite training provided to instructors.

Instructors using ILDs enjoyed engaging with the class, but spent significant time and energy adjusting to the innovation. Many found it difficult to stick to the provided format, and they struggled with logistical aspects of delivering ILDs. Having dedicated technical assistance helped alleviate some of these difficulties. Sharma et al. (2010) point out that contextual differences between the development and secondary site can affect student learning gains, but still found value in using the ILDs for student learning and instructor professional development.

**Implications for change:** A traditional approach to disseminating curriculum as a change strategy often involves education experts developing research-based strategies and curricular materials, collecting data on the value of the new approach, and sharing through journal articles, books, websites, talks, and workshops. This change model often assumes individuals will choose to change, and can take materials developed elsewhere and easily implement them “as is” and achieve similar success. However, secondary implementers often lack the funding, project team, course release time, and education experts that contributed to success at the development site. While Sharma et al. (2010) chose ILDs because they could work in large introductory courses taught in lecture halls, using the reform in their own context proved to be non-trivial. Even though physics education researchers helped with implementation, it took time for instructors to become comfortable using the ILDs, and their comfort using the reform seemed to affect students’ learning gains. Significant implementation difficulties could overwhelm potential users at other sites, for example, needing technical support staff and the up-front investment of time required for training and preparing for implementation (Sharma et al., 2010).

These findings indicate that improving communication between developers and secondary users can better facilitate educational transformation. Secondary users may benefit from knowing how to overcome common challenges, and developers need to document these difficulties. With enough support to get them through initial challenges, secondary users may experience the value of the innovation firsthand and be more likely to continue using it (Dancy, Henderson & Turpen, 2016). It is critical that education researchers study the challenges of secondary implementations so developers can support sites implementing RBIS in other contexts.

**Peer Discussion and Instruction**

Many instructors in physics and astronomy use classroom polling (Dufresne, Gerace, Leonard, Mestre & Wenk, 1996) to actively engage their students during class. Initially designed for use in lecture halls, students respond to a question presented by the instructor, discussing their answers with classmates before engaging in a class-wide discussion (Mazur, 1997; Crouch & Mazur, 2001; Prather, Rudolph & Brissenden, 2009). Instructors ask non-trivial, conceptual questions to target common misconceptions (Ding, Reay, Lee & Bao, 2009; Beatty, Gerace, Leonard & Dufresne, 2006), thus deriving real-time feedback on student understanding and inspiring interactive discussions. Peer instruction (PI) can be adapted to almost any course, class size and discipline.

**In the reform literature:** When Henderson and Dancy (2009) surveyed 722 physics
instructors across the United States, peer instruction was the most commonly known (by over 63% of people surveyed) and most often trialed (29% claimed to be current users) RBIS of the twenty-four in their survey. Yet self-described “users” rarely used peer instruction as intended, and may not know the developer’s recommendations. In fact, only 6.3% of respondents who said that they used Peer instruction “as described by the developer” in fact met all five criteria outlined by Mazur (1997). This shows how instructors either knowingly (or unknowingly) make extensive modifications during implementation (Henderson, 2005; Dancy, Henderson & Turpen, 2016). Occasionally, modifications are productive (Beichner, 2008), but other times, instructors lack the expertise of educational researchers and may modify techniques in unproductive ways. For example, instructors may undervalue the social aspect of learning (see, e.g., Redish, 2003) and drop the peer-to-peer interaction of PI (Henderson & Dancy, 2005), likely hampering its effectiveness.

Such differences between users’ conceptions and their enactment of PI also reveals that self-report does not reliably describe enacted practice in a way that aligns with researcher definitions. This is concerning since much of the reform literature depends on instructor self-report via surveys or interviews. Dancy, Henderson and Turpen (2016) interviewed 35 self-reported users or former users of peer instruction to further probe their use of this reform, and the decisions behind adoption and/or abandonment of this practice. More than half of those who recalled their first exposure learned about PI through an informal conversation with a colleague, echoing findings about the spread of JiTT (Gavrin, 2006) and SCALE-UP (Foote et al., 2014). But word-of-mouth conversations may lose details. Indeed, instructors tend to modify peer instruction during implementation and may unknowingly modify out critical elements, thus compromising effectiveness (Henderson & Dancy, 2009).

Turpen, Dancy and Henderson (2016) also used peer instruction as a case study to investigate perceived constraints and affordances that affect instructors’ decisions to use reforms. Many instructors want to avoid pure lecture but still teach in traditional spaces, so 20% of interviewees described choosing PI because it was easy to incorporate within a lecture environment—a low-risk, low-effort way to start reforming their courses. For this reason, PI is sometimes referred to as a “gateway reform,” a relatively easy innovation to experiment with, that can lead to use of further active learning strategies if it goes well (Turpen, Dancy & Henderson, 2016). Others chose to try PI because colleagues used it, or they felt that interactive teaching was valued in their department. However, the concerns of frequent users and nonusers did not overlap. Frequent users worried about maximizing productive student-to-student interactions, while intermediate users’ concerns reflect their own apprehension in navigating the use of PI or its features. Nonusers’ most common concerns span structural constraints, some outside their sphere of influence, including student shortcomings and conflicting personal commitments (Turpen, Dancy & Henderson 2016). Because different people have different concerns, professional development may need to address varied needs.

Such concerns may also lead to abandonment of the strategy. In general, Henderson and Dancy (2009) found high levels of disuse of RBIS, up to 80%, but Peer Instruction was toward the lower end of this abandonment scale (32%). Usually, instructors will keep using PI if they have positive experiences with it but abandon it if
they have difficulties (Dancy, Henderson & Turpen, 2016).

**Implications for change:** Because peer instruction is a low-risk “gateway reform,” it is important to support instructors through the trial period. Positive personal experiences were commonly mentioned as a reason to adopt or maintain peer instruction, while negative personal experiences were reasons to abandon PI. Instructors need to have strategies to overcome common initial difficulties; with such support, they will see value in the innovation and can gain confidence to try more radical reforms (Dancy, Henderson & Turpen, 2016).

Consciously or unconsciously, users rarely implement PI (and other reforms) as designed (Henderson & Dancy, 2009; Henderson, 2005; Dancy, Henderson & Turpen, 2016). Developers need to anticipate that their resources will be modified and adapted and should clearly identify what aspects of the reform are critical for success and the multiple ways to achieve these essential elements. PI may be considered a robust innovation that “degrades gracefully” under modification (Henderson, 2005, p. 785).

Finally, different user groups need different kinds of support (Dancy, Henderson & Turpen, 2016). Those considering the reform may need convincing of its value and advice about navigating structural difficulties. New users may need help navigating these difficulties until they learn how to modify it on their own. Current users may need very specific information about handling difficult student interactions, building questions that generate good discussions, and more. Professional development efforts should meet instructors where they are, which may require differentiated instruction to address the concerns of different user groups.

**TUTORIALS**

In physics, a tutorial, recitation, or discussion session often supplements the lecture giving students an opportunity to meet in smaller sections for two to three hours per week to apply course concepts. Typically, a graduate teaching assistant (TA) or undergraduate student will supervise a section of 15–30 students, to provide individualized help and feedback on conceptual activities or problem solving. The smaller class sizes promote deeper interactions between students and their instructors and provide a better environment for group work and collaborative learning. Usually reformed tutorials are accompanied by increased and improved TA training (Koenig, Endorf & Braun, 2007; Lawrenz, 1992).

Tutorials in introductory physics (TIP) were designed at the University of Washington (UW) to supplement the lecture in building understanding and logical thinking (McDermott & Shaffer, 1998). Researcher-designed activities elicit student ideas and preconceptions, create conflicts with observed phenomena, and use logic and self-consistency to resolve conflicts. Students work in groups on a series of pretests, worksheets, and homework assignments, while instructors play the role of “learning coach,” using Socratic strategies to ask guiding questions (Hake, 1992) without explicitly providing answers. A side benefit is that TAs often attend intense, weekly training sessions (Koenig, Endorf & Braun, 2007), which can lead to improvements in instructional practice for graduate students as well as to conceptual learning benefits for students. Extensions to the tutorial framework have incorporated concepts in mathematics and technological tools such as computers for data acquisition and displaying videos or simulations (Wittman, 2004), placed emphasis on developing both students’
concepts and epistemologies or beliefs about learning physics, and allowed instructors to modify the worksheets (Elby, 2001). In astronomy, lecture-tutorial curricula share many of the same principles (Prather et al., 2004). Tutorials have also been developed for upper-division electricity and magnetism courses (Chasteen & Pollock, 2008).

**In the reform literature:** Finkelstein and Pollock (2005) studied the adoption process of TIP at the University of Colorado (UC), Boulder to (1) examine whether secondary implementations could achieve gains comparable to the development site, and to (2) develop a “contextual constructivist” framework for understanding and organizing the features that shape successful adoption outside the development site. They documented “obvious and subtle” factors required for success, considering students’ engagement in individual tasks, the classroom situations in which these tasks are embedded, and broader departmental and educational contexts.

Finkelstein and Pollock (2005) found that the structural requirements for tutorials included functional space, modest equipment, well-trained staff, and aligned assessment. But ensuring that the norms and beliefs of the surrounding environment support the innovation may be even more difficult. For example, the department must value the tutorials enough to provide space, staffing (almost twice the typical number of TAs) and training. Instructors must establish a learning environment where students feel comfortable interacting with each other and with the instructor. Instructor expertise significantly impacts learning gains (Pollock & Finkelstein, 2008) and their beliefs about the nature of knowing can affect the classroom climate, either promoting or inhibiting productive interaction. Norms for students and TAs change too. For example, students must come to tutorials prepared with background knowledge and TAs must know how to ask guiding questions instead of disclosing answers.

**Implications for change:** Finkelstein and Pollock (2005) demonstrate the importance of subtle shifts in beliefs and culture that must accompany more obvious changes in materials and structures required for implementation. Establishing a classroom culture of interactivity, student-centered instruction and discovery-based learning was essential, they write:

*One of the main challenges of defining and producing replication remains: how does one adopt and adapt a culture of norms and beliefs? The materials and even structural implementation of tutorials are relatively easy to replicate; however the norms surrounding these cannot simply be stated and exported—rather, they must be grown and adopted. That is, while it might be useful to explicitly state appropriate norms, as we attempt here, faculty must adopt these and appropriate them to make sense at their institutions (p. 9).*

This work demonstrates it is possible for secondary sites to achieve comparable success to the original development site, if the secondary site’s norms and beliefs support the innovation. These results still should be interpreted cautiously. The University of Colorado, Boulder and the University of Washington have similar student populations, access to experts in physics education research, and grant funding facilitated this attempt at replication. Other institutions may have more difficulty achieving similarly high learning gains. Indeed, Pollock and Finkelstein (2008) found that even at the same institution, student performance varied by instructor background: classes taught by instructors informed by physics education research
consistently posted higher student learning gains than those of less-informed teachers. Either way, successful replication involves aligning tasks, norms, practices and infrastructure in the new context to support the innovation.

LABORATORIES

Most physicists agree that labs are a critical component of the physics curriculum, and many laboratory innovations have been introduced throughout the decades, including microcomputer-based labs (e.g., Sokoloff, Laws & Thornton, 2007), and labs focusing on the process of scientific thinking or on computational exercises (e.g, Chabay, 2015).

No one laboratory reform has been widely adopted at many institutions, because lab reform is notoriously difficult and resource-intensive with regards to equipment, space, and instructor time (Zwickl, Finkelstein, & Lewandowski, 2012). Many instructors agree on the goals of introductory courses, but labs can be used to achieve a variety of goals that are harder to articulate and prioritize, thus, the goals of lab instruction are often much broader than those associated with course lecture sections.

Compounding this lack of consensus on the aims of lab instruction, there is no dominant diagnostic instrument for assessing lab outcomes. Due to this variation in goals and curricula, choosing one representative example is difficult. However, Investigative Science Learning Environment (ISLE) has spread to universities and high schools around the country and inspired recent reinventions that embody contemporary trends, including IoLab (Selen, n.d.), making it an interesting case study. For more information on transforming labs, Zwickl, Finkelstein and Lewandowski (2012) describe the backward-design process of transforming CU Boulder’s upper division labs, with strategies on how to handle some of the unique challenges of lab reform.

Process-of-Science Labs

The Investigative Science Learning Environment (ISLE) (Etkina, Murphy & Zou, 2006) is a curriculum designed to help students replicate thinking processes closer to those of real scientists (Karelina & Etkina, 2007). Students meet in large rooms for interactive group activities involving complex problems and conceptual questions requiring multiple representations. In class, students go through a learning cycle of observation...
and exploration, developing explanations, designing experiments to test and refine their hypotheses and applying knowledge to solve problems. In ISLE labs, students may use traditional equipment but do not receive instructions on how to perform the experiments and need to design their own procedures. ISLE labs can be implemented without adopting the team meetings, so can be used in otherwise “traditional” courses (Demaree & Lin, 2006).

In the reform literature: ISLE was initially developed for small sections led by the course instructor. To extend ISLE to large classes, the developers reformatted guiding questions, provided self-assessment rubrics to aid in writing reports, and wrote reflection questions that emphasized process thinking. With appropriate TA training on managing students and grading reports, Etkin, Murphy and Zou (2006) found that ISLE labs could be used in large classes. Instructors at another large, research institution (Ansell & Selin, 2016) made further modifications, implementing ISLE-inspired labs in a way intended to reduce demands on classroom space, equipment, and instructor time. A pilot study provides evidence that ISLE methods can be used in a mixed at-home and in-class setting with versatile data collection devices.

Implications for change: Lab reform is difficult because the goals of lab instruction vary between institutions, and investment in expensive equipment adds extra inertia against change. Dancy and Henderson (2008) describe a reform continuum from adoption, where users are supposed to use curricular materials as-is, to reinvention, where users may be inspired by surface features but make significant changes to materials. User adaptations to their own contexts help to develop a sense of ownership that aids with sustainability (Rogers, 2010), but some modifications compromise the effectiveness of the reform. ISLE labs are a good example of a reform that has well-defined “essential elements”: users can adapt surface features to fit their own situation while maintaining its core integrity. The modular labs can be adopted with or without changing the surrounding course structure. ISLE developers explicitly modified the course activities to preserve “essential elements” in new settings, and adoptions indicate that these adaptations and hybrid models are successful (Demaree & Lin, 2006; Anell & Selin, 2016).

STRUCTURAL INTERVENTIONS

Some instructional interventions involve significant alterations to course structure, classroom environment, and scheduling. In Student-Centered Active Learning Environment with Upside-down Pedagogies (SCALE-UP), lecture, lab, and recitation are combined into a single interactive environment. While SCALE-UP reforms may take more effort to initiate, the physical space may play a role in increasing instructor interest (Foote et al., 2014) and sustainability of the reform (Knaub, Foote, Henderson, Dancy & Beichner, 2016; Foote, Knaub, Henderson, Dancy & Beichner, 2016). A structural reform in a more traditional space is Science One (Benbasat & Gass, 2002;
Dryden et al., 2012), an integrated first-year science program that was a challenge to get up and running but has been sustained for over two decades. Comprehensive reforms like these have the potential to improve student outcomes (Redish, 2003) more strongly than moderate alternatives.

**Studio Physics**

The term “studio” comes from the CUPLE physics studio at Rensselaer, a “comprehensive unified physics learning environment” (Wilson, 1994) that combines lecture and laboratory sessions into a single class session where 30-45 students meet to work collaboratively on computer-based activities. In the mid-1990s, this basic idea was adapted to large university physics courses (Beichner, 2008). SCALE-UP uses a very carefully designed classroom where students (typically 50–100) work at round tables in teams, with access to whiteboards on walls and multiple display screens around the room. Students engage in hands-on and heads-on activities as instructors circulate through the classroom, sparking Socratic-like dialogues and providing immediate, personalized feedback (Beichner, 2008).

SCALE-UP has inspired hundreds of secondary sites (Foote et al., 2014; Dori et al., 2003; Baepler, Walker & Driessen, 2014; Ingram, Jesse, Fleagle, Florman & Van Horne, 2013). Studies characterize “studio-like instruction” as using a reformed classroom space to facilitate face-to-face interaction, accompanied by pedagogical changes to promote active learning. Reforming the classroom space without changing pedagogy is not sufficient to increase learning gains (Cummings, Marx, Thornton & Kuhl, 1999), and some instructors still lecture in spaces with flexible seating (Stains et al., 2018). However, studio reforms can lead to improved conceptual learning, problem solving, increased attendance and retention, especially for members of underrepresented groups (Beichner, 2008).

**In the reform literature:** A web survey conducted in 2012-13 revealed that SCALE-UP style instruction was used at 314 departments in 189 institutions of higher education in 21 countries (Foote et al., 2014). As in studies of other RBIS, more respondents indicated learning about SCALE-UP via interpersonal channels—talks, workshops and colleagues—than via mass media channels. The authors estimate that dissemination of SCALE-UP in physics may be at the tipping point between adoption by adventurous early users and the more mainstream majority (see Figure 7.1). As in other studies, implementers demonstrate significant pedagogical and structural variation in their use of SCALE-UP.

In contrast to findings by Henderson et al. (2012) that about a third of people who use a RBIS abandon it, Knaub et al. (2016) did not find people who invested in a reformed classroom then discontinued use of SCALE-UP. Rather, if sites could assemble a group of faculty and administrators willing to invest in and construct a classroom, the implementation was likely to be sustained. Revising the classroom space sparked conversations about teaching that led to the formation of a guiding coalition who pushed the reform forward (Foote et al., 2016; Knaub et al., 2016 Kotter, 1995, 1996). The physical space encouraged instructors to interact with their students in new ways, leading to reformed behavior (Gaffney et al., 2010). Once instructors got comfortable using active learning in the studio space, they often brought active learning strategies into other classes they taught, even in traditional spaces (Knaub et al., 2016). The newly renovated, high-tech classrooms led to a sense of pride around innovative teaching. Some universities explicitly incentivized faculty to...
teach in the new space, and competition for this opportunity led to further validation. The classroom also attracted curious onlookers who asked how they could teach in the space, spreading interest organically, while campus tours to the new space highlighted progressive instruction as part of the institution’s identity (Knaub et al., 2016).

While the room certainly catalyzes positive changes toward innovative instruction, securing a classroom may require years of preparation (Fullan, 2000). Interviewees (Foote et al., 2016; Foote, Knaub, Henderson, Dancy & Beichner, 2018) described challenges that postponed reform until the right administrator took over, or funding or classroom space became available. Even if all the elements (Foote et al., 2016; Kotter, 1995) are not present at once, it is important for reformers to plant seeds by connecting with others in the organization—then when something shifts, implementation can proceed (Foote et al., 2018).

Implications for change: In dissemination of innovation theory, Rogers (2010) characterizes user groups by when they chose to adopt. As shown in Figure 7.1, Rogers describes the first adopters as innovators: social risk takers who are closely connected to scientific sources. Early adopters, the next group, are typically also educated, have high social status, and can strongly influence others. According to Rogers, innovators and early adopters both respond to the newness of an idea and are not deterred by things that may not work perfectly. However, the next wave of adopters, the early majority, are much more conservative and cautious, with different needs to be addressed before adopting. Foote et al. (2014) estimated that SCALE-UP may be approaching 16% adoption among physics departments, so the marketing strategy may need to change to appeal to the more hesitant early majority (Moore, 2002; Rogers, 2010).
It may be surprising that a more resource-intensive reform increases the chance of sustainability. However, the higher inputs required for the structural change provided enabling aspects that promoted implementation and eventual institutionalization (Knaub et al., 2016; Rogers, 2010). While SCALE-UP requires an upfront investment in the space, ongoing costs can be comparable, if not less than the traditional model, especially at colleges that do not use graduate students to run lab and tutorial sections. The reformed space catalyzes changes in behavior that can spread to other classes, attracts the interest of other innovative instructors, and helps universities advertise innovative teaching as part of the institutional identity (Knaub, et al., 2016).

**Integrated Science Program**

Since the requirement of a reformed space is distinctive to SCALE-UP reforms, how might this idea that “more radical may be better” apply to teaching strategies that do not involve reformed space? One example that has many parallels is the first year Science One Integrated Science Program at the University of British Columbia (UBC) (Benbasat & Gass, 2002; Dryden et al., 2012). While classes occur in traditional lecture halls, professors from biology, chemistry, mathematics and physics team-teach an integrated course that affects scheduling and costs twice per student over the traditional model. Despite administrative support, starting Science One involved significant challenges in overcoming disciplinary differences (Kezar & Elrod, 2012), securing funding and managing politics. However, once Science One was in place, the rotation of instructors caused reformed teaching to spread among the teaching team and into their disciplines. Moreover, enthusiastic engagement of instructors and students created a strong community around Science One and high visibility around campus. The interdisciplinary curriculum creates horizontal networking among departments that do not typically interact, gaining widespread support; involvement of the dean’s office means that vertical networking has helped to support the program as well. Thus, UBC has been able to sustain Science One for over two decades (Benbasat & Gass, 2002).

**In the reform literature:** While Science One is specific to UBC, Kezar and Elrod (2012) describe a nationwide initiative called Project Kaleidoscope to promote similar interdisciplinary teaching and learning environments nationwide. They used a three-stage model of institutionalization (Kezar & Lester, 2009) to plan change efforts on 28 campuses to create interdisciplinary courses. First, the system is prepared for change during mobilization, the change is introduced during implementation, and the system is stabilized in its changed state in institutionalization.

In the mobilization stage, the organization prepares for change by developing an awareness of the need for change, creating vision, galvanizing support through intensive and extensive discussion, and mobilizing leadership and collective action. Kezar and Elrod (2012) report similar difficulties to those noted by Benbasat and Gass (2002) when trying to encourage instructors to work together across departments with different cultures, structures and traditions. However, once those people are connected across the university, wider support means the reform effort will be less affected by administrative and faculty turnover. Implementation involves creating infrastructure and support for the reform, which may involve revised rewards and incentives, new facilities, additional resources or altered teaching loads. Often, offering interdisciplinary courses involved altering scheduling, changing the way credits are assigned and may impact instructors’
teaching style. While creating new structures is difficult, usually reverting these changes requires substantial effort so once the new infrastructure is in place, it may remain. Institutionalization is the final stage of the process, in which the innovations are incorporated into the value systems, culture, and day-to-day norms of the institution. Since most universities do not have an established way to recognize and evaluate interdisciplinary work, embedding these efforts often involves getting the wider university to value the program. Elements of this change process are seen in the case of Science One at UBC: as students progress into a wide variety of degree programs and instructors from multiple disciplines rotate through teaching, the large number of individuals influenced by the program helps transform the larger culture.

**Implications for change:** The cases of SCALE-UP and Science One exemplify that the resources and support required for more significant, structural changes can aid in a reform’s long-term sustainability, even if it takes more effort to get started. The wider buy-in required for these dramatic changes minimizes issues due to administrative and faculty turnover, and once new structures are created, reverting back to old habits sometimes involves more work than staying with the new system.

### INSTRUCTIONAL CHANGE REFORM

Up to this point, we have focused on instructional materials and methods developed and disseminated for course reform. To transform higher education, reform efforts must change teaching practice more generally, so this section describes initiatives aimed at modifying instructional practice to create more widespread change.

Henderson, Finkelstein and Beach (2010) reviewed almost 200 papers and grouped change efforts into four main categories: (1) disseminating curriculum and pedagogy; (2) developing reflective teaching; (3) enacting policy; and (4) developing shared vision (see Figure 7.2). Their characterization is based on whether the change effort targets individuals or groups, with either prescribed or emergent outcomes. We describe change strategies in physics that fall under each of these categories. Occasionally, strategies fall under multiple categories. Multifaceted change efforts may increase the chance of success because they focus on more than one aspect of the system simultaneously (Henderson, Finkelstein & Beach, 2010; Borrego & Henderson, 2014).

---

**DISSEMINATING CURRICULUM AND PEDAGOGY**

One type of strategy targets individuals with prescriptive outcomes, disseminating curriculum and pedagogy. In this category, change agents typically communicate their vision of good teaching to instructors, who are often encouraged to adopt the innovation “as is” (Henderson, Beach & Finkelstein, 2011).
Transmission-based Seminars and Workshops

Traditionally, many workshops fall into the prescriptive, individually focused “dissemination” category, where experts share specialized knowledge to tell others how to teach or organize their curricula, or maybe contextualize recommendations within a framework. While there are questions about how substantially such professional development efforts can affect faculty practice (Olmstead & Turpen, 2016; Seymour, 2002; Weimer & Lenze, 1991; Levinson-Rose & Menges, 1981), workshops are one of the most common methods to help instructors improve their teaching.

Since 1996, the American Association of Physics Teachers, the American Astronomical Society, and the American Physical Society have jointly presented the Workshop for New Physics and Astronomy Faculty (Workshop) to introduce new faculty to instructional ideas and materials based on physics and astronomy education research. Since the workshop has attracted approximately 25% of all new physics and astronomy faculty, it has the potential to massively impact teaching practice in the field as a whole. The Workshop consciously presents new faculty with a variety of subject-specific pedagogical options so that instructors can choose those that fit their needs. This is based on findings of Henderson and Dancy (2008) that when workshops or seminars focus exclusively on one strategy, faculty are often skeptical of reformers who seem to be “selling” a particular product under the claim that it will work “as is” in all contexts.

Henderson’s (2008) survey of workshop participants provides evidence that this workshop does expose faculty to new instructional strategies and many continue to expand their improved teaching practice years after the workshop, sometimes inspiring further changes by starting conversations in their departments. The Workshop presents a wide variety of options and motivates faculty to learn more, so under these conditions, it appears that one-time, transmission-based workshops can be successful with a well-defined target audience. However, presenting many strategies over a short period of time means faculty may not understand all the details underlying the strategies, including variations of strategies that may apply for different contexts (Henderson, 2008). Dokter (2008) studied two- or four-year college faculty attending an astronomy workshop and also found that short exposure could create significant changes in behavior. Heavy users of workshop techniques often took what they saw in the workshop, and modified or retrofitted it to fit their circumstances.

DEVELOPING REFLECTIVE TEACHERS

Approaches seeking emergent outcomes for individuals are described as “developing reflective teachers.” Change agents encourage instructors to use their own knowledge, experience and skill to improve their instructional practice with instructor-identified and -defined change outcomes (Henderson, Beach & Finkelstein, 2011). Teacher reflection may involve monitoring student thinking or student learning; comparing data with intended learning objectives; and making decisions to maintain, initiate, adjust, or terminate a teaching approach (McAlpine & Weston, 2000).

Reflection-based Professional Development

Since getting instructors to implement well-developed and innovative teaching ideas is difficult, taking a more reflection-based approach to professional development may align better with the research on faculty change. Two tools seek to help change
agents think about doing this: (1) Olmstead and Turpen’s (2016) real-time professional development tool to help presenters collect data and reflect on their workshops; and (2) Prather and Brissenden’s (2008) situated apprenticeship framework and examples of how to apply it to professional development in astronomy.

Olmstead and Turpen (2016) developed a real-time professional development tool (R-PDOT) to encourage workshop leaders to aim for more ambitious, emergent outcomes instead of just telling instructors what strategies exist. These include developing instructors’ abilities to notice student thinking, reflect on their instruction, and engage in collaborative discussions about instructional practice. The R-PDOT helps measure how interactive and prescriptive workshops are by capturing the focal points of faculty engagement. The authors encourage forward-looking activities that will help instructors bring instructional strategies into their classrooms such as collaborative analysis of student tasks and pedagogical approaches, and time to plan for future instruction.

Prather and Brissenden (2008) developed “situated apprenticeship” as a professional development strategy in astronomy that allows instructors to practice teaching strategies in an environment of peer review. Situated apprenticeship is “a learner-centered approach to professional development that purposefully engages instructors’ preexisting conceptual and pedagogical understandings of a particular instructional strategy and provides a pathway to improving both” (Prather & Brissenden, 2008, p. 4). Practically, workshop leaders present a strategy, such as Think Pair Share, and criticize an aspect of implementation together. Then they break into collaborative groups to develop their own implementations, practice, and then critique each other’s mock teaching experiences. This often elicits preexisting conceptual and pedagogical understandings, which might have impeded change but now can be resolved through discussion. Participant comments indicate increases in pedagogical content knowledge, skills and confidence that will allow them to be reflective users of RBIS (Prather & Brissenden, 2008).

**Paired Teaching**
The University of British Columbia (UBC) developed a paired teaching program (Stang et al., 2017) as an affordable way to provide long-term support for new faculty members adopting innovative teaching strategies. Two faculty members teach a semester-long course together, sharing all aspects of the course. A new faculty member is paired with an experienced reformed instructor so the semester becomes an extended professional development opportunity, with several advantages: (1) the design allows enough time for faculty members to change their beliefs and practices about teaching; (2) the experienced teacher can provide feedback and encourage reflection for the new instructor; and (3) the intervention is context-specific to a course of mutual interest.

Stang et al. (2017) find that new faculty used reformed teaching during the program, continued to use reformed methods after the program ended, and transferred strategies to new courses and contexts. The investment required in instructor time is relatively small, but initial evidence is so promising that UBC’s Faculty of Science gives every new faculty member the opportunity to participate in Paired Teaching.

**ENACTING POLICY**
Henderson, Beach and Finkelstein (2011) identified approaches at the intersection of organizational structures or environments and prescribed outcomes as “enacting
policy.” This involves developing appropriate environments, using rules, reward systems and support structures, to facilitate and motivate instructors to engage in specific or desired activities. When a change agent has a particular vision toward which they want instructors to work, it falls under this “prescribed” category.

**Large-scale Institutional Change**

The Science Education Initiative (SEI) aimed to create large-scale sustainable change by providing discipline-based experts to facilitate course transformation within departments (Chasteen, Wilcox, et al., 2015) at the University of British Columbia (UBC) and University of Colorado (CU), Boulder. This initiative tried to add support and structure that would sustain change efforts while aligning and coordinating the efforts of instructors and administrators across campus. Departments competed for funding to hire a Science Education Specialist; these specialists worked with faculty to transform courses: articulating learning goals and student difficulties, developing instructional materials, using conceptual assessments to monitor progress, then disseminating materials so future instructors would not need to start over (Chasteen, Perkins, Pollock, Beale & Wieman, 2011).

Chasteen, Wilcox, Caballero, Perkins, Pollock and Wieman (2015) reported that the initiative impacted about a third of undergraduate courses, over half the students, and almost half of the faculty in the funded departments at UBC. Among departments, success varied depending on the culture, organizational structure and chair, proposal timing, and the Science Education Specialist. Huber and Hutchings (2014) report high adoption of active-learning and supportive infrastructure changes in one department, but others did not progress as quickly and completely. They found that the Science Education Specialists were key to the reform, providing time, information and support that encouraged persistent use of the strategies. Most faculty supported by Science Education Specialists continued to use reformed methods during the time frame studied (Wieman, Deslauriers & Gilley, 2013).

Without the presence of a Science Education Specialist, sustainability varies. In CU Boulder’s upper division transformation for physics, two thirds of transformed courses continued to use half of the developed materials. However, with a “high fidelity cutoff” of using 70% of developed materials, the sustainability drops to 46% (Chasteen et al., 2015). Transferring materials in a well-organized course archives is difficult and affects the subsequent level of use (Chasteen, Wilcox, et al., 2015). The greatest threats to sustainability were high levels of faculty rotation through courses and reduced faculty ownership when Science Education Specialists and physics education researchers developed course materials (Chasteen, Wilcox, et al., 2015).

Chasteen, Wilcox and coauthors (2015) claim that the SEI proves that change is possible in university STEM departments without changes in the institutional incentive system. However, the SEI was expensive, costing $600–800,000 USD over five years at CU and $1.5–2 million CAD over six years at UBC (Wieman, Perkins & Gilbert, 2010). Without incentives, some faculty found it hard to engage and sustain changes in the face of research pressures, especially those at the pre-tenure and senior stages of their career. Sharing resources between instructors and developing an online depository was more difficult than anticipated, creating inefficiencies that limited sustained reform, especially in courses with multiple sections (Wieman et al., 2010).
TRESTLE (TRansforming Education, Stimulating Teaching and Learning Excellence) is a multi-institution effort to build on the SEI but propagate change with a smaller infusion of resources by building a community around course transformation (Bay View Alliance, 2017). Paired teaching allows departments to achieve some of the benefits of the Science Education Specialists, without hiring full-time personnel.

**DEVELOPING SHARED VISION**

Approaches at the intersection of emergent outcomes and environments are described as “developing shared vision” (Henderson, Beach & Finkelstein, 2011) that will support new modes of instruction. This can happen on the level of department, institutional unit, or institution. The change agent should empower individuals to come together to articulate then work toward collectively envisioned change.

**Tackling Undergraduate Equity and Improved Instruction Through Departmental Action Teams**

Following the SEI initiative, CU undertook another change initiative explicitly focused on systematically changing the culture to support change (Reinholz, Corbo, Dancy & Finkelstein, 2017). The change agent team articulated university-wide goals of supporting research-based, student-centered teaching and increasing equity and diversity. Then, within each department, a working Departmental Action Team (DAT) identifies issues and goals. Experts in educational research and institutional change facilitate joint meetings where DATs create shared vision, interpret data, strategize about departmental politics or meet with departmental leaders or outside experts, thus facilitating collective thinking about increasing equitable participation for STEM undergraduates. One team decided to collect data to inform a report and recommendations presented at a faculty meeting, which in turn led to creation of an equity committee with further capacity for change.

**INSTITUTIONS AND BEYOND**

There have been some recent multi-disciplinary efforts to transform teaching in multiple departments across some campuses, some supported by the National Science Foundation Widening Implementation & Demonstration of Evidence Based Reforms (WIDER) awards. This program encouraged applicants to employ widespread use of instructional materials and methods that have a convincing evidentiary basis of effectiveness, especially in high-enrollment introductory courses. Such efforts have targeted more than just physics departments to have far-reaching impact.
LEARNING ASSISTANT MODEL

CU Boulder’s Learning Assistant (LA) program is improving university teaching while symbiotically helping to recruit and prepare K-12 physics teachers (Otero, Pollock & Finkelstein, 2010). This has wide impact since improving K-12 teaching can improve the preparation and motivation of students prior to university. Similar to peer-led team learning (Cracolice & Deming, 2001), talented undergraduates are hired to facilitate small-group interactions in tutorials or transformed lecture, with support to improve their pedagogical content knowledge. This program helps achieve multiple goals: supporting faculty in teaching transformed courses, recruiting highly qualified future teachers, and engaging many faculty in transformed teaching and in teacher preparation. Finally, the program helps transform science departmental cultures to value research-based teaching as a legitimate activity for professors and students.

Otero, Pollock and Finkelstein (2010) report they increased the pool of well-qualified K–12 physics teachers by a factor of three compared to a small number before the initiative. Pollock and Finkelstein (2008) found the use of LAs helped promote sustained, high-fidelity use of the Tutorials with high learning gains. Endorsement by the Physics Teacher Education Coalition (Scherr, Plisch & Goertzen, 2014) has aided dissemination and supported implementation. At CU Boulder and other institutions (see e.g., Goertzen, Brewe, Kramer, Wells & Jones, 2011), the LA program is university-wide and benefits from having stakeholders across departments. A national network called the LA Alliance offers training materials, presentations and recommendations.

GRADUATE STUDENTS AS FUTURE FACULTY

Heller (2017) argues that the most efficient place to change higher education is where the fewest institutions input to the supply chain of future faculty: embedding training into graduate school. Half of all physics PhDs (potential future faculty) come from about 50 institutions (Mulvey & Nicholson, 2014), and graduate students have more flexibility for professional development than postdocs. Since most graduate students already serve as teaching assistants, using TA time to build skills and knowledge can lead to improvements in teaching without a large investment of time and resources. Often TA assignments are based on institutional need (Austin 2002a, b) without thinking about how to make it a valuable experience for graduate students. Programs should address TA assignments and preparation time, and may include department orientations, weekly teaching seminars, and mentoring (Heller, 2017; Lawrenz, 1992).

The Center for the Integration of Research, Teaching, and Learning (CIRTL) also focuses change efforts on graduate students who will soon fill faculty ranks (Austin et al., 2009). They train participants in a view of Teaching-As-Research to encourage instructors to use research to deliberately, systematically,
and reflectively implement more effective teaching practices, and they use Learning Communities to bring graduate students together across disciplines and experience levels for shared learning, discovery, and generation of knowledge. The original CIRTL program at University of Wisconsin, Madison involved graduate courses, intergenerational small-group programs, monthly round-table dinners, targeted workshops and an education-based research project (Pfund et al., 2012). Participating graduate students and postdocs reported that their experiences helped them adjust effectively and creatively to teaching demands in their new positions (Bouwma-Gearhart et al., 2007).

SUMMARY

The following provides an overview of the two questions driving this research and literary review in undergraduate physics and astronomy instruction, and identifies which domains are the most well developed with respect to reform, and those where more progress needs to be made.

1. What evidence from research and practice is available about the nature and extent of implementation of research-based reforms in STEM instruction? What evidence is missing or inadequate?

Within the last decade, researchers in physics education have realized that reformed curricula are not sufficient to change university teaching practice and have started to study secondary implementations and faculty change efforts. Many efforts have targeted introductory courses, where the curriculum is relatively standard between institutions, and accepted diagnostic instruments can be used to monitor the efficacy of changes. These courses reach large numbers of students and can affect whether a student decides to pursue STEM. At the same time, reforming first-year courses can be challenging, as curricula are overloaded and multiple instructors may have limited investment in the course. In contrast, for upper division courses, high faculty interest in majors’ courses leads to productive discussion and engagement around learning goals (Chasteen et al., 2015; for other examples, see Pollock et al., 2010; Chasteen et al., 2012; Jones, Madison & Wieman, 2015; Manogue et al., 2010). No publications describe how well these upper division resources transfer to secondary implementations.

While many of the published studies are from large research universities with physics education research (PER) faculty, there is far less information available about what is happening in liberal arts and community colleges, where smaller faculties, smaller class sizes and different research expectations may affect the transformation process. Larger departments may be more likely to include faculty with PER expertise as well. At CU Boulder, PER expertise led to extensive development of high quality materials and the use of validated assessments, which lowered faculty workload in transformed courses. However, the presence of PER experts limited the number of people involved in making changes and sometimes led to faculty feeling less invested in using the resulting materials (Chasteen et al., 2010).
et al., 2015). Thus, having PER expertise available may simultaneously enable and inhibit widespread change.

Henderson, Beach and Finkelstein (2011) point out additional shortcomings of current literature:

Most change agents belong to one of three isolated research communities (disciplinary-based education, faculty development or higher education research), each using a subset of available strategies; new work often does not build on prior empirical or theoretical work; and most published results claim success of the change strategy studied, but the evidence presented is often not strong (p. 977).

Facilitating conversations across these research communities could lead to sharing strategies to tackle current problems.

Moreover, much of the current literature fails to scientifically measure the impact of the change (Henderson, Beach & Finkelstein, 2011), relying so far on self-reporting of teaching practices, where faculty may claim to use a reform, but it is often different than the implementer intended. Studies where researchers observe classes for details of implementation have not been done in a uniform manner across multiple departments and universities. The isolated case studies reported so far can be hard to unify within a single theoretical framework. It is hard to connect “the degree of implementation” to the impact on student outcomes because of high contextual variation between courses, departments and institutions. Researchers and instructors may want data that disassemble the “essential elements” of reform and measures the impact on learning outcomes, but practically, this may be difficult to implement.

2. What changes have been observed or reported, in what domains? Which types of change seem to be more established, and which less so?

The most developed domain of reform is classroom practice and assessment. Many RBIS have been developed over decades and faculty awareness has increased, but the uptake and use of these strategies are significantly lower. The most popular changes seem to be low-risk innovations that can be used within a lecture setting, such as peer instruction. These innovations can have a high rate of disuse (Henderson & Dancy, 2009), especially if there are no wider departmental (Wieman et al., 2013) or structural changes that support the change. Structural reforms such as SCALE-UP (Foote et al., 2014; Knaub et al., 2016) require widespread buy-in and a monetary investment that seem to aid in sustainability. In the case of labs, structure can impede change. Many institutions have spent decades investing in lab equipment so are hesitant to make a major and costly change. Compounding this issue, the goals of lab instruction are not always clear (Zwickl, Finkelstein & Lewandowski, 2012), so labs are one of the most difficult components of a course to reform. Supporting faculty during initial reform challenges may lower rates of abandonment, but it can be resource-intensive and decrease faculty ownership of the change (Chasteen et al., 2015; Wieman, Deslauriers & Gilley, 2013).

Reform efforts in classroom practices and assessment are accompanied by a growing number of curricular resources, including textbooks informed by physics education research (i.e. Cummings, Laws, Redish & Cooney, 2004; Chabay, 2015). Online homework systems provide opportunities for students to practice with instantaneous feedback and reduce TA grading time, so
TAs can interact in the classroom (Bonham, Beicher & Deardoff, 2001). Using reformed instruction in large enrollment courses (like SCALE-UP (Beichner et al., 2008)) often increases the reliance on online homework systems; while these have improved over the years, many are expensive and have set question banks that are hard to modify. Attempts at open-source assignment systems have been made (e.g., LON-CAPA; Kortemeyer, Kashy, Benenson & Bauer, 2008) but high maintenance costs and fast-changing technology have impeded the widespread use of locally designed tools.

The second most developed domain is instructor development and preparation for teaching. The New Faculty Workshop for Physics and Astronomy (Henderson, 2008) has successfully increased awareness of many RBIS for new faculty. Efforts to develop learning assistant (Otero, Pollock & Finkelstein, 2010) and teaching assistant (Lawrenz, 1992; Holmes, Martinuk, Ives & Warren, 2013) training will also help prepare potential future faculty (Austin et al., 2009). The Physics Teacher Education Coalition (PhysTEC) has created sustainable programs aimed at improving the quantity and quality of high school physics teachers (Scherr et al., 2014). Department-based efforts to better prepare physics teachers often simultaneously transform university teaching through course reform, Learning Assistant programs, outreach opportunities and more.

The next most developed domain is instructor interest, beliefs, and attitudes related to reform-based teaching. Change agents are increasingly aware that unidirectional dissemination of research-based materials in a one-off presentation rarely changes faculty behavior; some examples of other models exist (Reinholz et al., 2017) that seek to create cultural change to support reformed interest, beliefs and attitudes. While change agents increasingly recognize it is important to change instructor beliefs and attitudes in order to change their practice (Dancy & Henderson, 2010; Henderson, Dancy & Niewiadomska-Bugaj, 2012; Henderson, 2005), there are not widely accepted tools to measure and monitor instructor attitudes in higher education. Yerushalmi, Henderson, Heller, Heller and Kuo (2007) have made a preliminary attempt to map instructors’ beliefs about introductory physics teaching, so curriculum developers can design materials accordingly.

The least developed domain relates to diversity and inclusion in undergraduate STEM instruction. Some professional development efforts seek to increase instructors’ awareness of equity issues (Austin et al., 2009). Other examples in physics focus on student experiences, such as the COMPASS program at UC Berkeley (Albanna et al., 2013) and the COMPASS-inspired Access Network. These six university-based programs work with graduate and undergraduate students across the country towards a vision of a more diverse, equitable, inclusive, and accessible STEM community (The Access Network, n.d.). Overall, however, there has not been much published about widespread, multi-institution efforts to address equity in instruction.
There is evidence that interactive instruction may disproportionally benefit students from underrepresented groups (Beichner et al., 2008; Brewe et al., 2010; Lorenzo, Crouch & Mazur, 2006; Rodriguez, Potvin & Kramer, 2016). General course reform efforts may indirectly improve diversity by increasing classroom interaction, but this is not sufficient to remedy the complicated combination of factors that leads to inequity in participation and performance (Pollock, Finkelstein & Kost, 2007). McCullough’s (2018) review of research on gender and under-represented ethnicities in physics education research describes broadening research themes throughout the decades, writing that “the new research has pushed the boundaries beyond performance gaps and curricular issues, looking into bias, attitudes, affirmation and sense of belonging, and other people-focused issues” (p. 8).

Although that connection is not explicit, change agents could use this expanding literature to design initiatives to support increased equity.

FUTURE WORK

Funding agencies increasingly require applicants to explicitly discuss dissemination and sustainability. This contributes to an increased awareness of these aspects of the change process and more publications about change efforts in STEM education. While there may not be consensus for the best strategy for creating systematic change in university physics, there is evidence that two commonly used change strategies do not work: (1) developing and testing “best practice” curricular materials, disseminating them through scholarly avenues and expecting faculty to use them as-is; and (2) “top-down” policy-making meant to influence instructional practices without attention to situational constraints, instructor beliefs or how to support faculty through initial challenges (Henderson, Beach & Finkelstein, 2011). Reform efforts are context-dependent, and changing a system requires understanding the system before designing a compatible strategy.

Successful change involves aligning the change with the beliefs of individuals involved or finding ways to change beliefs so that alignment is achieved. Usually substantial change in beliefs requires extended interventions that support faculty through initial challenges, ideally accompanied by institutional or structural changes to institutionalize the transformation.

Hopefully, change will get easier as students who went through reformed courses work as teaching assistants in graduate school, then start to occupy faculty ranks. Physics has a strong and growing community around education research, where reformed instruction is expanding, but there are still challenges in the reward system and institutional structures that make widespread acceptance difficult.

REFERENCES


Elby, A. (2001). Helping physics students learn how to learn. American Journal of Physics, 69(S1), S54-S64.


Huber, M. T., & Hutchings, P. (2014). The Carl Wieman Science Education Initiative in the Department of Earth, Ocean, and Atmospheric Science, University of British Columbia. (Unpublished benchmark case study, Bay View Alliance).


LEVERS FOR CHANGE: An assessment of progress on changing STEM instruction
INTRODUCTION

The disciplinary reviews (Chapters 2-7) prepared as part of our work provide detailed and thoughtful analyses of the available evidence about the extent to which research-based reforms have been adopted by STEM instructors, and what has influenced these changes. The reviewers’ work also primed rich cross-disciplinary conversations among the broader group of meeting attendees. In this chapter, we summarize findings from both scholarly and practitioner assessments of these questions:

▲ What is the current state of research-based reform in undergraduate instruction within these six clusters of STEM disciplines?

▲ How did each arrive there? What levers for change—activities, events, influences, movements, groups, documents, contexts—have been important in reaching this state? And how are these levers similar or different by discipline?

▲ What provides evidence for these trajectories of change, and why?

▲ What can be learned from this evidence about how to expand and deepen the impact of these changes in the next decade?
This chapter begins with a cross-disciplinary perspective on themes of the six reviews. Then we offer a synthetic analysis of the working groups’ findings about the state of change, the sources of leverage on that change, and promising next steps.

**REVIEWS OF STEM EDUCATION RESEARCH**

In this section we summarize the key findings from the six scholarly reviews, highlighting commonalities and contrasting differences. The reviewers unearthed a variety of studies that measure the extent of uptake of active and student-centered teaching methods. Several of the studies indicate that 15-25% of STEM instructors are using active learning approaches extensively, and a similar proportion are using these approaches to some degree. Lecture is still the predominant teaching mode, with studies of varied samples indicating 50-75% of instructors using lecture as their main teaching method. Some samples show lower percentages of active learning. For example, low use of active-learning approaches in calculus (Johnson, Ch. 6) may reflect particular challenges of implementing these approaches in large enrollment courses that are often taught by teaching assistants. In most cases, the results cannot be compared directly across studies, because they vary in how they defined and bounded types of instruction, in the samples and time periods they investigated, and in the research methods they used. Nonetheless, the general consistency of these measurements is striking and provides a useful benchmark for RBIS uptake.

The studies also show interesting differences by discipline, most tellingly revealed in
large cross-discipline studies including both classroom observations and surveys. In a large classroom observation study, Stains and colleagues (2018) found that the proportion of classes observed to use lecture-heavy, “didactic” teaching styles was 40-65%, with interesting disciplinary differences. “Relative to chance, mathematics and geology have more student-centered styles than expected, biology has more interactive styles than expected, and chemistry has more didactic styles than expected” (p. 1469).

Drawing on data from surveys conducted by UCLA’s Higher Education Research Institute, Eagan (2016) reported that teaching in all the STEM disciplines showed less use of various student-centered teaching methods than seen in the arts and humanities, but the use of “extensive lecture” in STEM classes has declined in the decade 2004-2014 to percentages in the 60-70% range. He also found that STEM faculty infrequently report using “techniques to create an inclusive classroom,” consistent with the reviewers’ reports that this is an area of poor progress across the STEM disciplines. While the observation and survey studies categorized disciplines differently, mathematics is one field where findings can be compared. In the observation study, active-learning approaches are relatively abundant in mathematics, and this is corroborated by time trends in the HERI data that show significant growth in active and inquiry learning in mathematics.

Scale may be one influence on the spread of research-based teaching approaches within a field. McConnell (Ch. 5) argues that the prevalence of active-learning approaches in geosciences is due in part to the penetration of research-based materials and professional development into a smaller instructional workforce. However, scale does not alone explain the apparent higher degrees of uptake in large fields such as biology and mathematics, or the lower levels seen in chemistry (see Table 8.1).

The reviews note other points of general agreement in the literature across disciplines. Reform in assessment practices lags behind reform in other teaching practices. The reviewers noted continuing emphasis on memorization and content and on curve-based grading in most fields. Likewise, research-based strategies for laboratory instruction are less well developed than active engagement strategies for the lecture hall. This is ironic, given the natural learning opportunities afforded by laboratory investigations, but may reflect substantial structural challenges for changing laboratory instruction (see discussion in Foote, Ch. 7). In general, the body of education research that could support planning for lab reforms, as well as the research tools and methods needed to study laboratory learning and experiences, are also less developed. A related issue is that, in most fields, preparation of graduate students for current and future teaching roles—whether in laboratories, recitations, tutoring centers, or leading their own sections—is not yet a salient activity within the discipline. However, some reform projects have taken specific advantage of greater nimbleness within recitation sections, such as Tutorials in Physics, Peer-Led Team Learning, and activities guided by undergraduate Learning

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Biology</th>
<th>Chemistry</th>
<th>Engineering</th>
<th>Geosciences</th>
<th>Mathematics</th>
<th>Physics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of positions</td>
<td>49,180</td>
<td>21,080</td>
<td>37,310</td>
<td>10,690</td>
<td>50,630</td>
<td>13,540</td>
</tr>
</tbody>
</table>

Table 8.1: Number of STEM instructors at US college, universities, professional schools and junior colleges
(Data from US Bureau of Labor Statistics, May 2017; prepared by David McConnell)
Assistants. Still others have dropped the distinction between lecture and lab by blending experiments, simulations, and sense-making activities in a ‘studio’ format, as in Workshop Physics.

Some studies highlighted by the reviewers offer important insights about individual change processes. Overall, the affordances an instructor experiences from using active learning may be much more impactful than the barriers, thus, more attention should be paid to these affordances in talking with instructors about RBIS (Andrews, Ch. 2). Sustaining the use of RBIS is a bigger challenge than building awareness or encouraging a first try, so we need to learn more from persistent adopters (Finelli & Borrego, Ch. 4; Foote, Ch. 7). And there is more to learn about how instructors who are more experienced or more expert in active learning think, believe, and work, how they analyze student discourse on the fly in class, and make use of information from assessments.

The reviewers pointed out several methodological challenges in this kind of research. Teaching is a complex activity and it is difficult to characterize teaching reliably (AAAS, 2013). Researchers need clearly defined terms to classify instructional practices, but practitioners may interpret these terms differently. Instructional styles are personal, and the very act of categorizing them for research may be off-putting to instructors. For example, it was previously useful to refer to RBIS as “innovative,” but this terminology is now dated, and its continued use risks reifying RBIS as edgy or risky, when the data here show that they are becoming mainstream. Studies of fidelity of implementation are challenging to conduct but raise important questions. Student reports of their learning experiences are under-used but may be powerful in identifying essential elements of RBIS, corroborating or contradicting faculty reports of classroom practice. It is also challenging to track progress at a national level. It may be important to pay closer attention to differences by discipline or by unit of analysis (Johnson, Ch. 6; Stains, Ch. 3).

The chance to gather baseline data is past, and it can be hard to recognize the origins of change, pinpoint diverse influences along the way, and acknowledge changes in context such as funders’ increasing expectations that projects contribute to the knowledge base. Foundational projects such as the NSF-funded calculus reform efforts, engineering coalitions, and chemistry systemic initiatives have clearly been important in paving the way even if it is not possible to attribute specific outcomes to individual projects. The rise of DBER scholarship as a field has changed the landscape of STEM education reform, but insights from DBER work do not automatically penetrate into practice. The cases of geoscience (McConnell, Ch. 5) and physics/astronomy (Foote, Ch. 7) offer interesting contrasts here: physics education research is seen as essential in spurring change in physics and astronomy, but in geoscience education, research has followed rather than driven change.

Finally, the reviewers identified gaps in current knowledge about research-based instruction in STEM higher education, including:

▲ Large-scale studies that include representative samples of multiple types of institutions;

▲ Investigations of how RBIS are being implemented that address quality, not just quantity or extent of RBIS use, including studies that identify the essential elements of instructional approaches against which practice can be benchmarked;
Examination of RBIS in all elements of a course, including laboratory instruction, assessment, and out-of-class learning, and of RBIS use throughout a program or major;

Understanding of how teacher knowledge, beliefs, and attitudes contribute to effective and sustained use of RBIS, and how these may differ by discipline;

Rigorous investigations of whether and how professional development influences instructional practices, for whom, in what contexts, and over what time frame; and

Evaluation of whether and how instructors, instructor preparation efforts, and departments are attending to equity and inclusion in teaching, and what strategies are effective in strengthening instructors’ awareness and ability to create fully inclusive environments.

KEY FINDINGS

Chapter 1 describes the May 2018 meeting where we asked each disciplinary expert group to discuss the state of uptake of research-based instructional strategies in their STEM field. Expert groups were asked to provide an overall assessment of progress, as seen from their varied vantage points as change leaders and scholars working in many arenas. Their knowledge is distinct from the research results presented by the reviewers (Chapters 2-7) and summarized in the first section of this chapter, but it complements those findings and suggests directions for further research. In this section we summarize the working groups’ assessment of the state of change in their field, describing each field individually to provide some nuance on how perceptions from various disciplines may be similar or different. Quotations are taken from collaborative documents and posters where each group captured and shared their discussions.

The working group in life sciences recognized solid progress in adoption of RBIS: “We’re better than we used to be, but we’re not done yet.” For this field, Vision & Change in Undergraduate Biology Education (AAAS, 2011) has been “a rallying point,” a “point of nucleation” that is the “start of a movement” because it has brought together practitioners from across biological disciplines and communities to agree on core concepts for students to learn. They expressed a sense that use of RBIS in the life sciences was crossing a threshold or tipping point, seeing the glass as “more than half full.” Some of the challenges they see ahead are related to the interaction of RBIS with “the tyranny of content” coverage: What is essential to teach given the wide span of the field, and how should instructors and departments make decisions about what to teach? Other challenges include preparing instructors to use evidence-based teaching and valuing their work on teaching.
The working group in chemistry and biochemistry characterized progress as “moderate.” Instructors are aware of active learning and think they should be doing it, but often see their choice between RBIS and lecture as a stark one, all or nothing, a view that may inhibit some from trying smaller increments of change. To build on chemists’ epistemology of research, it may be important to foster a culture of trying and documenting “teaching experiments.” Assessment practices and laboratory instruction lag behind classroom instruction, though some progress from “cookbook” to guided-inquiry labs was reported. Overall, they noted, “There are multiple areas of progress, little peaks and eruptions everywhere, like the Hawaiian Islands,” reflecting many but uncoordinated efforts. They imagined “a field of flowers,” each flourishing to different degrees, depending on the nourishment offered by institutional, department, disciplinary, and individual components of the environment.

In engineering and computer science, the working group likewise saw a mix of bright spots and needs for further progress. “We’ve made progress on awareness of RBIS, but mostly lecture prevails. There are pockets of good practice, but our field still needs more widespread adoption and continued use.” Most progress has been made in courses for first-year students and senior capstone experiences, while the middle years of the curriculum are still content-heavy and pedagogy-poor. This group imagined influences on change through an atmospheric metaphor: high-flying “shooting stars” of well-adopted practices; lower-floating “clouds” that represent encouraging progress but are not yet widespread; and “rocketship” practices that are potential high flyers but “still haven’t taken off.” Connections between engineering/CS education research and instructor practice were imagined as “a rope bridge, when we need the Golden Gate Bridge.” The gap can be crossed, but the pathways are not robust. Like others, they noted lack of coherence in change efforts: “We could use a more coherent blueprint for change that really guides the entire community.”
The working group in geosciences viewed the field as “progressing,” and other groups saw progress in geoscience as enviable. Often spread across multiple departments within any one institution, the rally point for geoscience instructors has not been in departments or societies, but through instructor communities built around On the Cutting Edge (https://serc.carleton.edu/NAGTWorkshops/index.html) and related projects. This model has been powerful in combining professional development with the production of learning materials that are shared and used by others. Geoscience education research has been helpful but is not seen as the driver of change. Some aspects of how geoscience is situated within the STEM landscape were seen as affordances:

*Our intro classes are not required anywhere, which means we can experiment, and have to experiment—we must be entrepreneurial. Figuring out ways to broaden our field has driven a press for societal relevance, and a focus on interest and engagement. We have focused on different entry points to geoscience careers... as well as intellectual entry points... as ways to bring new ideas in education and geoscience research to the community. ...Our discipline and our departments are not required to exist within universities..., so we must pay attention ourselves to the health and efficacy of our structures, communities and curricula.*

By the same token, however, they noted, “We struggle to fully operationalize the value proposition for geoscience careers and paths, and personal impact on the world and society, relative to other fields. We don’t make a strong case for why students should come and stay in geoscience.” Thus, the perceived urgency of change has been important, as well as specific mechanisms that gather instructors for professional development and community-building.

In the mathematical sciences, the working group described a sense of optimism around the current pace of uptake of RBIS. “The interest is there; this is an opportune moment.” Flexibility has been important: “We are not stuck on one approach to active learning. There are existence proofs in many kinds of institutions.” Examples of flexible approaches include the growing instructor community around inquiry learning (inquirybasedlearning.org) and the SEMINAL project involving graduate institutions in undergraduate mathematics reform efforts (http://www.aplu.org/projects-and-initiatives/stem-education/seminal/index.html). With increased attention to pathways
across the transition between high school and higher education come possibilities for synergy and collective learning about common challenges, such as adopting RBIS in large classes or coordinated multi-section courses, and to address strong (real or perceived) expectations of topical coverage from client disciplines.

The group contrasted the role of education research in this field (known as RUME, research on undergraduate mathematics education) with what they saw as more practical, instruction-focused research in other DBER fields:

*Our studies have grown out of K-12 research... [sometimes focusing narrowly on things like] talking with two students about how they think about limits. We tend to forefront theory and methods, not instructional change. ...We haven’t been in conversation with the faculty leading professional development for faculty."

The epistemology of mathematics affects how math educators receive or view education research:

In physics and biology, one driver of change has been large-scale quantitative studies based on data, evidence and claims. In math, we work with theorems and proofs, so that affects what we think counts as evidence—sometimes one compelling story can be all that is needed.

They noted the advantage of pre-existing, multiple connections among the professional societies in mathematics and statistics. Compared to some fields, “we do better at cooperating in all kinds of ways.” In statistics education, a collaborative community and strong professional society leadership have similarly fueled uptake of RBIS in that discipline.

In physics and astronomy, the working group recognized that much progress has occurred through development of specific RBIS that, in many cases, embed active learning pedagogies in specific curricular materials. Here, physics education research (PER) has been “exceptional in advancing” the menu of choices, particularly for lower-division courses. Yet the linkage of active
instructional methods to materials also heightens issues of fidelity and adaptation, and whether instructors use the RBIS in ways that generate the maximal student outcomes—“probably not.” Thus, as in other fields, the bottleneck to faculty uptake was not the development of materials, but instructors’ discovery, adoption and effective use of these materials. And like other groups, this group identified laboratory instruction and assessment as areas that lag behind classroom RBIS.

Across the disciplinary groups, these assessments identify similar areas of progress and challenge, while acknowledging differences in the paths that disciplines have traveled. Meeting participants concurred broadly with the disciplinary reviewers’ findings that progress in instructor uptake of RBIS has been notable and substantial. While the proportion of instructors who do not use RBIS still exceeds the proportion who do, “the center of gravity has shifted.”

However, participants and reviewers also concurred that attention to classroom equity and inclusion in instruction is inadequate and insufficiently persuasive. There is broad awareness of the evidence about unequal outcomes for STEM students, and some projects have begun to incorporate equity into their work with instructors, but there is not yet a widely felt sense of urgency nor sufficient action to infuse and integrate equity considerations into efforts in instructional development, nor sufficient information about the varied experiences of RBIS for students from different groups. While research shows active learning strategies can help to close gaps in student outcomes and experiences by leveling the playing field, these approaches do not magically or automatically improve equity (see Killpack & Melón, 2016, and literature reviewed in Laursen & Rasmussen, 2019). Thus, it is not sufficient for instructors to adopt RBIS; they must be alert, knowledgeable and skilled in attending to classroom dynamics and student experiences as they learn and apply RBIS. Deeper understanding and better practical strategies will be needed to engage instructors, especially those from majority groups, in these aspects of research-based STEM instruction.

The next section describes in some detail the levers for change that participants identified and analyzed in their discussions as helping to shift the center of gravity. Following that, Section 4 focuses on “promising arenas” for future work to enable higher-order approaches to expand and deepen the impact of research-based instruction and address critical issues of equity and inclusion.

EXPERT WORKING GROUPS’ ANALYSES OF LEVERS FOR CHANGE

As discussed in Chapter 1, participants in the working meeting were asked to draw
on their observations and experiences as a knowledge base for considering progress on STEM instructors’ uptake of RBIS. In this chapter, “Key Findings” summarizes their collective assessments, by discipline, and their general sense that progress has been made, and compares these views to those of the disciplinary reviewers, summarized in the first section, “Reviews of STEM Education Research.”

The expert groups were also asked to discuss how we had arrived at the present state: what levers for change had been important so far, including activities, events, influences, movements, groups, documents, or contexts, and a series of structured discussions led them through this topic. From the groups’ initial, open-ended discussions, we distilled a list of 15 levers and asked each group to review them systematically. Members of each expert group worked together to provide a joint assessment of the extent to which each lever had been influential in generating instructional change in their own field, rating each as having high, medium, low, or no influence. They supported these assessments by identifying how the lever had operated in their field, and why it was important or not.

The analysis in this section is based on careful qualitative analysis of themes and patterns in the groups’ ratings and their justifications for those ratings. During that analysis, two levers were collapsed into others, due to the similarity with which they were viewed by the groups. The disciplinary groups’ assessments of the final thirteen levers for change are categorized and summarized in Table 8.2.

Note: Unless described otherwise, the quotations below are taken from the expert groups’ documentation of their discussions, or occasionally from transcribed meeting sessions. They are not attributed to specific people or working groups except where it is essential to differentiate perspectives from different disciplines.

LEVERAGE ON MOTIVATIONS FOR CHANGE

Four of the levers act mainly to provide motivation or a rationale for change, i.e., identifying the need, suggesting solutions, and prompting individual instructors or their departments and institutions to adopt RBIS. These levers may provide initial impetus to change, by posing a need or requirement for change, but they may also support change that has begun through other means by providing arguments that instructors and leaders can use to defend or justify their work, or that change agents can use to argue for funding or recruit participation.

Accreditation and Certification of Departments, Programs, or Institutions

Note: Initially, we asked groups to consider disciplinary accreditation processes separately from accountability processes originating elsewhere, such as expectations from state legislatures that institutions surpass or improve upon particular measures of student success. However, working groups treated these levers very similarly, so they are combined here.

Across disciplines, discipline-based accreditation processes were seen as having relatively low influence on change in instruction. They focused on curriculum, resources, and teaching loads, not pedagogy. From the perspective of disciplines who do not use such processes, accreditation processes were seen as “stifling,” “constraining” or requiring “fit to a particular form,” and thus their absence enables experimentation, innovation, and customization to local contexts. Chemists and engineers, whose fields do have
discipline-based certifications, felt these processes did accommodate innovation, but agreed that the perception of constraints from accrediting bodies was often used as an excuse not to change anything. For engineering, ABET’s introduction of outcomes-based accreditation in 2000 was initially a high-influence lever that focused attention on student outcomes, but its importance has declined in the decades since then. In practice, the group noted, “The ABET process promotes innovation rather than stifles it,” but it is often cited as a barrier.

In chemistry, departments may elect to offer a major that is certified by the American Chemical Society (ACS); while the ACS guidelines for certification do call for “pedagogies that have been shown to be effective,” no measurements are specified and no consequences are attached. Because the ACS guidelines are perceived to encourage breadth over depth of curriculum, they may discourage methods that are thought to take more time and offer less content coverage. Many departments use the ACS’s standardized exams for specific

<table>
<thead>
<tr>
<th>Primary Function</th>
<th>Primary Level of Operation</th>
<th>Lever</th>
<th>Life sciences</th>
<th>Chem &amp; biochem</th>
<th>Engr &amp; CS</th>
<th>Geosciences</th>
<th>Math &amp; statistics</th>
<th>Physics &amp; astronomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motivate</td>
<td>All</td>
<td>Accreditation or certification of departments, programs or institutions by disciplines or states</td>
<td>Low/med</td>
<td>Low/med</td>
<td>Low/med</td>
<td>Low/med</td>
<td>Low/med</td>
<td>None</td>
</tr>
<tr>
<td>Motivate</td>
<td>All</td>
<td>Guiding documents from professional societies or other leadership bodies</td>
<td>Med/high</td>
<td>Low</td>
<td>Low</td>
<td>Med</td>
<td>Med/high</td>
<td>Low/med</td>
</tr>
<tr>
<td>Motivate</td>
<td>All</td>
<td>Demands from employers for specific competencies among graduates</td>
<td>None/low</td>
<td>Low/med</td>
<td>High</td>
<td>Med</td>
<td>Low/med</td>
<td>Med</td>
</tr>
<tr>
<td>Motivate &amp; Prepare</td>
<td>All</td>
<td>Results from education research</td>
<td>Med/high</td>
<td>Med/high</td>
<td>Med/ high</td>
<td>Med</td>
<td>Med</td>
<td>Med</td>
</tr>
<tr>
<td>Prepare</td>
<td>Individuals</td>
<td>Professional development of instructors, graduate TAs, future faculty</td>
<td>Med/high</td>
<td>Med/ high</td>
<td>Med</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Prepare</td>
<td>Individuals</td>
<td>Resource collections or digital libraries of instructional materials</td>
<td>Low/med</td>
<td>Low/med</td>
<td>Low</td>
<td>High</td>
<td>Low/med</td>
<td>Med</td>
</tr>
<tr>
<td>Prepare</td>
<td>Individuals</td>
<td>Educational technologies to facilitate teaching and learning</td>
<td>Med</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low/med</td>
<td>Low/med</td>
</tr>
<tr>
<td>Prepare</td>
<td>Individuals</td>
<td>Communities of practice built around specific pedagogical practices, courses, or sub-disciplines</td>
<td>Med</td>
<td>High</td>
<td>Med</td>
<td>High</td>
<td>Med/ high</td>
<td>Med</td>
</tr>
<tr>
<td>Motivate</td>
<td>Academic Units</td>
<td>Models and exemplars from other institutions</td>
<td>Low/med</td>
<td>Low</td>
<td>Med/high</td>
<td>High</td>
<td>Low/med</td>
<td>Med</td>
</tr>
<tr>
<td>Motivate</td>
<td>Academic Units</td>
<td>Local data and evidence about student outcomes</td>
<td>Low/med</td>
<td>Low/med</td>
<td>Low</td>
<td>Med</td>
<td>Med</td>
<td>Med</td>
</tr>
<tr>
<td>Prepare</td>
<td>Academic Units</td>
<td>Collaboration with other disciplines or departments</td>
<td>Low/med</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low/med</td>
<td>High</td>
</tr>
<tr>
<td>Prepare</td>
<td>Academic Units</td>
<td>Local leaders and internal change agents with a vision</td>
<td>High</td>
<td>Low/med</td>
<td>Med/high</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Stimulate</td>
<td>System</td>
<td>Federal and private funders investments in STEM education</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

Table 8.2: Levers and levels of change.
chemistry courses (e.g., General Chemistry) that are perceived as a benchmark of educational quality. Changes to these exams could thus be a tool for helping faculty think about how they prepare students. Moreover, because of this benchmarking, student scores on an ACS test can be useful to demonstrate that “no harm is done” when students “cover” less but learn more deeply (Gutwill-Wise, 2001). Accreditors’ requirements for student success data can also be a tool for departments to acquire such data.

With respect to evidence, one group noted, “Funding agencies have driven the need for evidence, rather than accrediting bodies.” However, program review could function as a lever on instruction, for example when review processes established explicit expectations for program outcomes and evidence that these are being met, and thus led departments to think more carefully about curriculum and assessment. For this to occur, local leaders need to “align and connect the dots” to relate the ultimate goals set by legislatures or accreditors (such as graduation rates) to more proximal measures (such as passing rates in gateway courses) to education practices in those courses (such as instruction and student support). In the process, disciplinary resources, such as guiding documents and comparative data could become helpful.

Some groups saw potential in a different type of accreditation that measures valued outcomes, such as student diversity, learning, or success rates, perhaps offering a highly visible endorsement, as in the model used by the AAAS’s SEA Change project (https://seachange.aaas.org/). Institutional program review connected to higher education accreditation processes sometimes serves this function for individual departments.

Finally, the leverage offered by accreditation and certification processes is generally different for two-year colleges (2YCs). Courses and programs in these institutions must be aligned with state four-year (4Y) systems to enable student to transfer courses and credits. Thus, pressure from legislatures and other state bodies could be influential, particularly in mathematics, where lack of success in developmental courses can throttle students’ pursuit of a degree in many fields. This lever operated through both educational policy and funding structures, as the mathematical sciences group noted:

*Moves to performance-based funding, combined with policy determining placement and math requirements (and the attendant demands for related evidence), have dramatically shifted institutional policies and structures, and hence priorities and allocation of resources. In some cases, this has led to more equitable placement, clearer articulation, and redesign of course content, structure, and instructional approach. (emphasis in original)*

Many 2YC STEM programs also go through additional accreditation processes to be able to offer industry-specific professional certifications from industry groups. These processes combine to add layers of oversight and work for instructors and administrators. Further, meeting multiple sets of requirements may constrain instructors’ choices around curriculum and pedagogy. This “extra emphasis for community college faculty may thus present a barrier to change (if there’s conflict between the proposed change and the bureaucratic processes) or could be leveraged to support change (if the proposed changes are synergistic with the existing processes).”
Guiding Documents from Professional Societies or Other Leadership Bodies

Guiding documents and statements from professional organizations were generally viewed as having moderate to low influence on instructional change. Such documents were helpful in providing common ground or starting places for individual departments, multi-campus or state systems. Some disciplines have found that these documents had more traction than did others, for example:

▲ In the life sciences, *Vision & Change* has been an important influence (see Andrews, Ch. 2; also Austin, 2018). *Vision & Change* is seen to have had greater impact than prior such reports, perhaps because of the strong efforts made to gather people to build consensus and discuss recommendations.

▲ In mathematics, the unanimity and strength of recent statements about active learning issued jointly by multiple societies has “provided cover for those who are advocating for change in their departments—a critical resource in the current climate, particularly in two-year college contexts.”

▲ In statistics, these documents have been powerful: the smaller number of statistics educators enables consensus documents to more readily penetrate and move the field, and demand from client disciplines or professional fields who receive these students may generate more discussion and foster consensus about important ideas and methods.

Several groups noted that disciplinary society recommendations about curriculum are more common and carry more weight than those from non-disciplinary sources. The case of mathematics offers an interesting twist, because recent documents focus on pedagogy and not just curriculum (e.g., CBMS, 2016; GAISE, 2016; MAA, 2018; Saxe & Braddy, 2015), and are being used and cited widely at professional meetings, workshops, and journals. Results from a recent nationwide research study of the characteristics of effective calculus programs has similar reach (Bressoud, Mesa & Rasmussen, 2015). “These position statements help to get local buy-in for instructional change,” the math/stats working group noted. Because guidance on the mathematics curriculum from the professional society is a long-standing tradition (Steen, n.d.), there may be habits of awareness and receptivity that applies as well to pedagogy guidance.

Also, from mathematics, there is evidence for a different type of impact from a national report. By calling out mathematics preparation as an urgent challenge and suggesting that it may be necessary for other disciplines to take over teaching college mathematics to their own majors, the PCAST report *Engage to Excel* (2012) galvanized mathematicians into a vigorous response (Bressoud, 2013; Bressoud, Friedlander & Levermore, 2013; MAA, 2012; White, 2014) that is now having positive effects. Overall, guiding documents are most powerful when they reflect consensus across multiple leadership bodies, and when they are not merely documents but processes of engagement, gathering people and generating conversation at all stages and leading to active dissemination and professional development efforts. This is easier when disciplinary communities already are convening: “In chemistry, we have the opportunity to reach the whole discipline, because all chemists are at the same meeting.” In contrast, because it is a large and diverse field without a single disciplinary gathering point, the engineering/CS group...
thought that a process like that used for development of Vision & Change might be useful in their field.

In some fields, data from professional societies have been useful, providing data that local leaders can use to make comparisons and drive arguments for department-based change. For example, in physics and astronomy, the American Institute of Physics, an umbrella organization serving multiple physical science societies, has a well-respected statistical arm that surveys departments and compiles data on enrollments, degrees, and diversity. In geoscience, the American Geosciences Institute offers similar information.

Demands from Employers for Specific Competencies Among Graduates

The influence of this lever was perceived as low by some disciplinary clusters, but medium-to-high in fields with a more applied orientation, including statistics and engineering/CS. In geoscience and chemistry, experiences of observing graduates’ difficulty in finding jobs were impactful for some departments locally, although these was not seen as pushing instruction on a broader scale. “Most faculty don’t focus on this, but they may respond to the pressures these demands create in their institution’s administration,” noted another group. In chemistry, such pressures might lead to incorporation of specific skills, such as laboratory safety, but did not encourage use of RBIS. In biology, the feedback loop from employers was viewed as particularly weak, because so many life science graduates do not work in the field:

Examples of how employer expectations had influenced instruction in engineering included needs for graduates with skills in teamwork and design, which had led to curricula and instruction designed to cultivate these skills. Similar potential was seen for “supporting our students to be agents of change around diversity and inclusion to influence the environments they enter.”

Education Research Results

This lever was seen as having medium influence across disciplines. Discipline-based education research (DBER) is “effective within our community [of DBER scholars and well-informed educators] but has little impact beyond,” noted one group. “It raises awareness but is not driving practice,” agreed another. “It helps to have this as part of the argument, but results are often too general to translate into practice or local policy,” noted a third. “Anecdote and personal experience tend to be more compelling” than research. We asked the disciplinary discussion groups to separately consider DBER and basic research on how people learn as levers, but most did not find this a compelling distinction. While DBER scholars recognized basic research on cognition as useful underpinning for their own work on disciplinary learning, for classroom instructors, such work was “too abstracted” from actionable findings:

Faculty seem to have little patience for basic research in the learning sciences, perhaps because it seems too far afield from everyday practice, and the current ways of disseminating this research are not conducive to translation into everyday practice.
Some noted that research in social psychology was now reaching instructors in useful ways, in areas that influence student learning such as stereotype threat and growth mindsets. Such research “suggests that there are other approaches than personal experience,” which helps instructors accept the findings.

Despite these caveats, education research was seen as an influential lever under certain conditions, particularly when the “preponderance of evidence” became large enough to compile into reviews or meta-analyses (e.g., Freeman, et al., 2014; Hake, 1998; Ruiz-Primo, et al., 2011). In these cases, instructors may even be persuaded by education research results from studies done in STEM fields other than their own. For example, the widely cited meta-analysis by Freeman and coauthors seems to be compelling across disciplines, even among those not well represented in that data set, such as geoscience. Similarly, studies with large samples were more impactful: “‘Big N’ research seems to have greater impact than smaller case studies or qualitative studies.”

Instructors at teaching-focused institutions—primarily undergraduate institutions (PUIs) and two-year colleges—or in teaching-focused roles were thought to be more likely to read education research than were STEM research-focused faculty at research institutions. Faculty change agents who could lead and encourage change in their disciplines or departments were a particularly important group to reach with education research results. For all audiences, education research results were most impactful when presented in “accessible, ready-to-implement ways” offering “practical and actionable strategies that are easy to adopt by ‘garden-variety’ faculty” and that can be adapted to diverse educational settings. This is not to suggest that instructors will adopt curriculum wholesale, merely because it is research-based, but rather speaks to the potential for education research as a lever of another sort, when findings are communicated in ways that yield practical advice on what and how to change, a step beyond results that provide reasons to change.

Groups agreed that communicating existing results “with a marketing mindset” was important for making research practically useful for educators, but there was little consensus on who should be responsible for this type of communication. DBER scholars are typically better prepared to do this work, but many are reluctant to have their work seen as practical, applied or popular because they have fought hard to gain credibility for education research as rigorous scholarship in their disciplinary departments. Institutional reward structures favor the scholarship of discovery over scholarship of integration or application (Boyer, 1990), as one group noted:

Discipline-based education researchers’ research is removed from the faculty members trying to implement teaching. They view it as their research and scholarship, what they stand on as a faculty member. They don’t always believe their job is institutional improvement.

Another group noted, “Implementing best practices can have a big impact on the campus but doesn’t lead to big publications.” In some fields, the National Academies’ DBER report was seen as helping to legitimize DBER work with STEM disciplinary colleagues (Singer, Nielsen, & Schweingruber, 2012). Such a shift aids DBER scholars’ credibility among STEM colleagues but does not necessarily enable them to take the step toward making their work accessible to practitioners.

Some argued that, rather than researchers, well-informed practitioners were best placed
to communicate useful research findings to other educators, and indeed this likely occurs informally. This communication also happened through embedded education specialists (who may or may not be DBER scholars) in disciplinary departments (see studies of Science Faculty with Education Specialties, SFES, by Bush and coauthors, referenced in Andrews, Ch. 2), local projects in scholarship of teaching and learning (SoTL), and researcher-practitioner collaborations or mutual mentorship. Teaching and learning center specialists are a strong asset on some campuses and underused on others. Some fields have conferences where researchers and practitioners mingled readily, such as the Earth Educators’ Rendezvous in geoscience. A community effort to synthesize findings from geoscience education research recognizes “the necessity for the community to collaborate, to more fully understand why and how people are learning” and seeks to identify grand challenges relevant to research and practice that can positively affect teaching and learning about the Earth (St. John, 2018).

In chemistry, the Biennial Conference on Chemical Education targets practitioners, but has included more research in recent years, and the Journal of Chemical Education offers both research articles and practical articles on pedagogical strategies or laboratory experiments. Several groups suggested that new and different forms and venues are needed to strengthen research-practitioner communication. For example, Finelli and Borrego (Ch. 4) suggest, based on the wide reach of the meta-analysis by Freeman and coauthors, that systematic literature reviews may be a helpful tool for distilling large bodies of research and helping non-experts identify key themes across the research.

In sum, education research is a lever that operates in two categories. As a lever for motivating or justifying change, education research results that demonstrate the impact of RBIS help to shift the discourse, particularly when results mount up to the level of reviews and meta-analyses. It is also clear that research findings alone do not persuade instructors to change. And, as a lever for showing instructors what and how to change, education research results may be made more powerful than they have been so far. As one group put it:

*When geared toward target audiences, education research results can have strong impact. The community would benefit from focusing on communicating existing results with a marketing mindset. This is the place where there is a large gap between research and practice.*

That is, educational research acts as a lever for change when it is translated and taken up by instructors and institutions, which is by no means automatic. To realize this promise, it may be necessary to cultivate a STEM higher education specialty equivalent to the growing field of translational research in medicine. Steps toward this might include developing new forms and venues for communicating results and their implications for practice, investing in professional development to foster expertise in this type of communication, and identifying disciplinary and institutional structures and rewards structures that enable and encourage people to specialize in this type of communication. To meaningfully
inform collective improvement efforts, such work must necessarily attend to a wider swath of education research, not just DBER work. For example, work that is important for STEM education reform on topics such as growth mindsets, equitable teaching practices, and models for organizational change (Laursen & Rasmussen, 2019) may come from other subdisciplines of education and from other social sciences.

**LEVERAGE ON INDIVIDUAL INSTRUCTORS**

Four more levers (Table 8.2) can best be described as offering mechanisms for change. Once instructors are motivated to change their teaching to use more RBIS, what helps them to actually do so? Mechanisms can also precede motivation. For example, based on decades of study of K-12 teacher professional development, Guskey (2002) suggests that professional development need not first persuade instructors to change their minds before they implement a new strategy. Rather, he proposes, the goal should be to get them to try something new, and they will be persuaded when they see an impact on their students:

> ...It is not the professional development per se, but the experience of successful implementation that changes teachers’ attitudes and beliefs. They believe it works because they have seen it work, and that experience shapes their attitudes and beliefs (pp. 383-384).

Thus, this discussion of levers does not presume a model of how and when different levers are applied but includes examples of different ways they may be influential. In this section, we discuss four prominent levers for change that support individuals to change their own teaching. Later, we discuss four additional levers that apply to change of instructional practice across departments and institutions. As before, quotations are taken from the expert working groups’ written discussion records and transcribed discussions.

**Professional Development for Instructors**

Across disciplines, professional development (PD) about teaching was seen as a lever of high importance. A good PD “experience can provide opportunities to change perspective and practice. These experiences can be very powerful and can normalize ways of thinking about teaching” that are new to the participant but are supported by research and used by others. In this way, PD can lead to culture change in a department or disciplinary community formed by people who teach a certain course or use a certain teaching approach. “It’s difficult to envision how active learning scales [up] without opportunities for instructors to experience new models for instruction,” noted one group.

One problem with assessing PD as a lever of change is the shortage of research evidence for its benefits. The review chapters describe specific PD programs that have been studied, but also note many challenges to such studies. There is a need for better measures of teaching that make it possible to characterize change in practice as a result of PD (AAAS, 2013), and long timescales are required to enact and discern meaningful change in teaching practice, much less to document changes in student outcomes (see also Giersch & McMartin, 2014). This problem applies to research studies but also for instructors who wish to track their own
outcomes over time. Overall, the groups noted, “The jury is still out on PD. ...Its impact will vary widely on the depth and quality of the professional learning experiences offered and taken up.”

One exception to this pattern came from the geosciences, where strong PD programs have been sustained over many years, coupled with materials development and sharing as a way to engage instructors in thinking about learning goals, activities, and assessment. McConnell (Ch. 5) summarizes compelling evidence for the impact of this model of professional development. But in engineering, strong programs such as the National Effective Teaching Institute have not penetrated the field sufficiently to generate widespread change, despite strong evidence for their positive impact on those who did participate (e.g., Felder & Brent, 2010).

To be effective, professional development must be of high quality and offered widely. PD opportunities through the discipline were not available to everyone, but PD opportunities on campus could be uneven in goals and quality. PD opportunities that create sustained community have been more likely to catch hold and have potential to change culture. For example, in physics, the New Faculty Workshop and, in astronomy, workshops led by the Center for Astronomy Education (CAE) have reached thousands of instructors, especially newer instructors. As these instructors become aware of effective teaching practices, they become “ambassadors to their departments—a critical scale-up step where change happens.” While examples of effective virtual communities were rare, one example is the long-lived CAE opt-in email listserv, AstroLrner, which connects over 1000 current practitioners in a virtual “support group” for astronomy educators (Slater, n.d.). The experience of AstroLrner suggests that a tight focus on lower-division astronomy courses and very active facilitation, prompting and moderation are crucial for supporting and sustaining such a community. Hayward and Laursen (2018) have recently used social network analysis to identify social and practical mechanisms that make a listserv supportive as instructors implement RBIS.

Professional development opportunities for regular college instructors with faculty appointments were viewed as more readily available than those for supporting graduate teaching assistants (GTAs) in their instructional roles or in preparing future faculty. In most fields, a lack of networks for TA training means that PD for GTAs may be very uneven in quality and availability. “We are good at teaching TAs how to teach [students] how to use a microscope and lead field trips, but we are not yet strong on teaching them how to use active learning,” noted the life sciences group. The Center for the Integration of Research, Teaching and Learning (CIRTL Network) offers professional development to future faculty—STEM graduate students and postdocs interested in academic career—on its ~40 member campuses and through online courses (www.cirtl.net ). Recent studies document CIRTL participants’ self-reported gains in their knowledge, skills, and confidence around teaching, and how these gains help when they start a teaching career (Connolly, Lee, Savoy & Hill, 2016; Connolly, Lee & Savoy, 2018; Pfund, et al., 2012; Prevost, Vergara, Urban-Lurain, Campa, 2018; Schein, et al., 2017; Vergara, et al., 2017).

Recent projects have begun to address the need for professional development of those who can lead instructional change. One example is the Partnership for Undergraduate Life Science Education (PULSE), a community that has developed tools and strategies for improving their own abilities to lead change and applied them in supporting
life science departments across the US in implementing recommendations of Vision & Change (http://www.pulse-community.org/), thus amplifying their own effect as they work with a wider circle of departments. In mathematics, the Academy of Inquiry Based Learning (AIBL) is leading a major initiative to provide professional development workshops on inquiry-based learning for college mathematics educators at the same time as it trains workshop leaders and builds professional development leadership capacity in undergraduate mathematics education (http://www.inquirybasedlearning.org/). The College Mathematics Instructor Development Source (CoMInDS) supports people who lead professional development for GTAs in mathematics departments, rather than targeting the GTAs themselves (cominds.maa.org). It will take some time before the discipline-wide impact of such strategies can be observed or documented.

In sum, professional development is seen as a powerful lever for change because, as college STEM instructors, “we were never taught to teach.” Effective PD models will support individuals’ growth in teaching knowledge, skills and reflection and “do not see faculty knowledge, readiness and motivation through a deficit model,” and they must attend to instructor identity in relation to teaching. Reaching graduate students with professional development on teaching is seen as a strong lever for the future.

Communities of Practice
Closely related to professional development in both importance and approach is the formation of communities of practice (CoP) that invite instructors to join a group gathered around a particular pedagogical practice, STEM topic, course, or sub-discipline. Some long-lived examples of these include

- The POGIL Project, which originated in chemistry and now supports Process Oriented Guided Inquiry Learning in several STEM disciplines (www.pogil.org)
- InTeGrate, organized around sustainability topics in introductory geoscience courses (https://serc.carleton.edu/integrate/index.html)
- SENCER, Science Education for New Civic Engagements and Responsibilities, emphasizing general education science courses that connect core science concepts to complex, real-world issues and societal challenges (http://sencer.net/)
- IONic, the Interactive Online Network of Inorganic Chemists, connecting people in a specific subdiscipline of chemistry who may teach both lower- and upper-division courses (https://www.ionicviper.org).

These examples—and many others that were offered—illustrate a range of approaches to how a CoP may be conceived. For example, POGIL is explicitly focused on supporting guided inquiry as a pedagogical approach that improves student learning of key concepts and scientific practices, while IONic emphasizes development of learning materials and assessment of student outcomes, with less explicit focus on the use of research-based pedagogies. InTeGrate is disciplinary and SENCER is interdisciplinary. These long-lived examples also highlight how continued support from funders—in some cases across more than two decades—has been essential to sustain each CoP’s work and grow the community. Other challenges for CoPs include sharing leadership and advocacy roles, developing future leaders, staying true to their mission and maintaining cohesion while remaining open to new people and new ideas, and finding the right balance between face-to-face and virtual activities (Haberler, Laursen & Hayward, 2018; Kezar & Gehrke, 2015; Kezar, Gehrke & Bernstein-Sierra, 2017; Kezar & Lester, 2009). Some found the general
absence of students from these communities to be a limitation. And relatively few CoPs have made the leap from grant support to self-sustaining as a nonprofit organization; one exception is the community around Peer-Led Team Learning (Dreyfuss, 2013).

What CoPs offer as a lever for change is “the means to share practice, knowledge, and expertise among people who have an affinity for a particular topic. A virtual CoP can help adopters or those in the field to receive guidance and find resources that can help with adoption.” Through workshops, materials, and access to other practitioners, CoPs can be used to:

▲ Foster exchange of lessons learned and good practices, making teaching more visible (Reisner & Williams, 2010) and strengthening teaching identities;
▲ Generate new understanding of how to develop and implement a promising practice;
▲ Offer collaboration and accountability to support continued development and use of the practice;
▲ Verify the effectiveness and benefit of practice, including transferability;
▲ Validate, publish and disseminate specific practices;
▲ Accelerate decision-making and implementation of best practice;
▲ Achieve higher standards in projects, strategies, and improving outcomes; and
▲ Engage leading experts and develop future leaders.

Overall, meeting participants viewed CoPs as a promising but not fully vetted lever for change—a “growth area” for many disciplines:

There have been some strong examples of how cross-institutional CoPs can promote improvement of practice, spread of innovation, and sustainability of change. Local communities, particularly when supported by the institution in a sustained way, have also provided foundations for change in culture and practice. However, CoPs (particularly across institutions) are notoriously difficult to grow to critical mass for self-sustainability.

CoPs may be especially influential when they couple other levers together, such as in geoscience where professional development has been linked to materials development (McConnell, Ch. 5), or when professional development provides a route to pull people into an ongoing CoP, as seen for inquiry-based learning in mathematics (Hayward, Kogan & Laursen, 2016; Hayward & Laursen, 2018). There is much yet to learn about how to build and sustain robust and supportive CoPs.

Resource Collections and Digital Libraries

Resource collections were seen as having low or medium-low influence as a lever of change. Collections of vetted materials are appealing because they lower the start-up costs of a new instructional approach and provide images or “existence proofs” of effective lessons. And good activities can have very high influence, if they are adapted locally by groups of instructors teaching a common course.

However, the existence of good materials is alone not sufficient to generate change, as instructors must find them, trust them, and adapt and use them. “There are almost too many resources, too spread out, and unwieldy to browse.” Indeed, an analysis of NSF’s efforts to establish digital libraries in STEM fields noted that the libraries had been unable to keep up and compete with the Internet at large (Mervis, 2009). In engineering, “we learned early on that overly prescriptive curriculum materials don’t get adopted.” Likewise, in physics, research by Henderson and colleagues has demonstrated that developing and disseminating research-
based educational materials is alone insufficient to encourage faculty to use RBIS in a lasting way (Foote, Ch. 7).

Cases where digital resource collections were most helpful often focused on smaller topics or communities, rather than trying to be comprehensive. In statistics education, the Consortium for the Advancement of Undergraduate Statistics Education is a significant resource that offers workshops, webinars, data sets, curricular materials and assessment tools (www.causeweb.org). In mathematics, the Journal of Inquiry-Based Learning uses an open-access journal model for sharing IBL problem sequences for specific courses, thus serving a specific community of practice (http://www.jiblm.org/). Similarly, in biology, CourseSource seeks to overcome some of the barriers faced by prior digital resource efforts (www.coursesource.org). Peer review of contributions provides a professional incentive for instructors to share their materials, and the use of learning goals set by professional societies offer a way to organize materials around common learning goals and a route to engage disciplinary societies in educational leadership. And, as discussed above, in the geosciences, development of materials and assessments was effective because it was coupled to professional development and engagement in a community of practice (McConnell, Ch. 5).

What was effective was not just materials, but “thoughtful grouping and curation of existing numerous teaching resources, tied together with specific instructional and local learning goals and outcomes.”

Educational Technologies

Educational technologies were rated low or medium-low as a lever of change. They were seen as useful tools to incorporate appropriately into any teaching approach, not an end in and of themselves. Some saw data from digitally mediated education as providing “an opportunity that has not been fully realized,” including special opportunities for research. But:

Technology is not a magic pill. It should follow, not lead. Instructors’ philosophy, motivation and mindset need to change first. If technology can help that, great! But technology should not come first—too often, it does.

In mathematics, classroom technologies are useful in fields that rely on data and visualization, such as statistics and applied math, but nearly invisible in proof-based mathematics. In engineering, the CATME system helps with the practicalities of creating and assessing student teams (info.catme.org ). Gathering answers to thought-provoking conceptual questions using clickers or other classroom response systems can serve as ‘gateway’ practices to engage instructors and students in active learning (Stains, Ch. 3; Foote, Ch. 7). And implementation of flipped classrooms may be simplified by technologies to support independent pre-class work via videos,
podcasts, readings, simulations or practice problems. Online homework is used in many chemistry and physics classrooms, but there is not yet good evidence for whether it supports teaching strategies or improves formative feedback to students.

Finally, one group noted the importance of including fully online and hybrid courses, and the people who teach them, in thinking about uptake of RBIS: “[Online teaching] is growing. How do we ensure that reform reaches these courses and instructors?” There may be a need for studies of instructor knowledge and practices to inform professional development on best practices in online teaching, and for studies that examine how established RBIS developed for face-to-face settings can be adapted for online environments (e.g., J. Smith, et al., 2014).

LEVERAGE ON DEPARTMENTAL AND INSTITUTIONAL CHANGE

Four additional levers (Table 8.2) operate at a different level. Rather than motivating or enabling instructors to change individually, they operate on groups of instructors that are organized in departments, schools or institutions. The levers identified earlier can operate on groups, too – for example when instructors from one department adopt new materials or attend professional development together – but the levers discussed here explicitly operate at a different level of the system, the academic unit. Two of these levers are largely ones that work to motivate or justify change, and two are mechanisms or processes by which change can occur.

Models and Exemplars from Other Institutions

The disciplinary groups varied quite widely on how they viewed the importance of models and exemplars from other institutions as a lever for change. Those seeing this lever as high in influence cited “peer pressure as a positive force and noted the value of “war stories” and “lessons learned.” They described strong communities of practice, for example in statistics education, as providing the venue for exchange of ideas. Exchanging speakers among institutions is one of the ways we define scientific disciplines, noted one group, and “having some of these exchanges focus on effective education normalizes this work as part of a discipline.” Another group observed that DBER positions had been created in some departments because “people saw other institutions doing this.”

In engineering, several recent winners of the Gordon Prize awarded by the National Academy of Engineering represent programs built around RBIS such as experiential and problem- and project-based learning. A recent report on the state of the art in engineering education (Graham, 2018) also highlights the leadership of programs based in RBIS such as Olin College and Iron Range Engineering, and NSF funding support has elevated exemplars of student success models from the alliances supported by Broadening Participation in Computing program. The visibility of such programs in prestigious company supports the engineering/CS group’s view that, in their field, “new, innovative programs provide peer pressure and models for making change in existing programs.”

In contrast, those arguing that this lever was low in influence described the “not invented here” mentality. For example, “it’s usual to say it won’t work in our context,” as one group noted. Differences in real or perceived availability of resources is another reason that departments do not look to others for inspiration or ideas. “We don’t have aspirational institutions for teaching” as we do in research, noted the chemistry group. Others acknowledged truth in both views, stating, “It depends on how your institution
Local Data and Evidence About Student Outcomes

Local data on student outcomes were seen as having low-to-medium influence currently, but as a lever with potential for more significant impact. “If this were a significant driver, people would be mortified! This should drive change in our discipline,” noted one group. Some knew of departments that had been motivated to address high DFW rates (i.e., drop, fail, withdraw) or improve retention. And data of other types could help units “think about demographics, recognize the supports and barriers faced by students (e.g., families, socio-cultural structures), meet the students where they are.” Student data should measure not just learning but also belonging and persistence: “Inclusion should be a big piece of the evidence on student outcomes, and it isn’t yet, not the way it should be.”

Groups identified some conditions under which local data could be useful. First, they pointed out that “what is considered evidence depends on the environment, but it is important to have common measures.” For local data to be influential, it should be focused within single departments. Comparison to peer departments in other places could be useful because “it eliminates a common objection that ‘our situation is different.’” For example, in physics, Force Concept Inventory gain scores (Hake, 1998) collected from a wide array of institutions prompted many to consider moving to active engagement methods. However, overly aggregated data, such as across all departments in a college of engineering, could obscure useful patterns, and data comparing units within an institution could be used for elitist arguments rather than for improvement. “It’s important to know who is collecting the data and the role those individuals play,” noted one group; involving all stakeholders in considering data was more successful than top-down efforts.

Among participants with some experience using institutional data as a lever of change, these data were seen as effective particularly in persuading administrators. What’s needed as “evidence for reform within a class does not have to be what is convincing as a measure of institutional change. Sometimes a department chair does not care about a pre-/post-test within the class, but they want to know about success in the next class, or DFW rates.” One meeting participant with experience in multi-institution change efforts noted:

We have found local evidence about student outcomes to be the primary sources of evidence that are meaningful to institutional leadership. Large N studies of effectiveness (particularly with respect to downstream student outcomes, such as retention, graduation, transfer, etc.) may also be attended to, but not as powerfully.

For example, the Carnegie Math Pathways (https://www.carnegiemathpathways.org/) has used educational analytics to good advantage to inform educators about their students, document improvements and challenges, and persuade administrators of the urgency of problems and effectiveness of solutions.

The groups were in broad consensus about the potential for local data as a more powerful tool, but they acknowledged that its use is not yet well developed or understood. “We don’t have a coherent model for using such data,” noted one group. Said another, “We’re not sure how to make this work yet—we can collect and share data, but we don’t know how to make it make a difference.”
Others noted difficulties in extracting local data from their administration or institutional research office.

To realize the promise of data as a lever for change, it will be important to identify useful measures, analytical methods, and visualizations, to experiment with platforms for exploring and reporting data, and to cultivate a spirit of transparency around data both within and across institutions. A recent study (NASEM, 2018a) has addressed this problem at the national level, but locally useful tools and processes can be shared collegially within and across disciplines much sooner than a national system will be available. For example, the geoscience group wished for “a toolbox of assessment instruments that could be used across the community—an assessment bank, so to speak. [It could show us] how to promote more effective assessment for what we want—and when we find out we are not meeting [our objectives], then we are motivated to change.” Such tools could be used as formative assessment to identify needs, challenges and improvements, but would also help to set community standards.

**Collaboration with Other Disciplines or Departments**

Collaboration across disciplines or departments was seen as having modest influence by most of the disciplinary groups. Instructors can be “very siloed, even within departments,” noted one group. For example, in chemistry, teaching responsibilities for particular courses are strongly held by subdisciplines, in contrast to physics where instructors are expected to teach a wide range of courses within the field. Instructors may also prioritize time and travel funds to share their work at disciplinary conferences, rather than attending multi-disciplinary conferences across the sciences or across undergraduate education. As the engineering/CS group noted, “We tend to reinvent the wheel a lot, even when we should look to what already exists,” yet there is “immense opportunity” and need to connect with arts and humanities (Gunckel & Tolbert, 2018; NASEM, 2018b).

Some participants had experience working with other departments. Those with large student enrollment, such as biology in many institutions, recognized they had some leverage on other units that taught their students, such as mathematics and chemistry. But in mathematics, where courses serve students from many disciplines, diverse and often conflicting expectations from other departments have often been problematic for prioritizing pedagogy and making curricula coherent and conceptual. At the same time, however, the math/stats group recognized opportunities for impact in connecting with other schools in their region, for example to address shared concerns across the high school-college transition for equity and student success in early college mathematics courses that serve as gateways or barriers to students’ education and career goals.

A few participants who had experience with significant interdisciplinary collaborations recognized significant potential in reaching beyond individual departments within institutions. They felt it was critical to have a sizeable number of people engaged in the vision and recognized the need to support people who were involved, such as paying lecturers and community college faculty who invested time in collaborative work. “Extending these discussions beyond STEM disciplines can help us to see the invisible constraints on our approaches and consider alternatives. We have implicit assumptions in our field about what has to happen, such as high-stakes, end-of-semester exams, but we could look to other disciplines...”

---

*Source: NASEM (2018a, 2018b)
LEVERS FOR CHANGE: An assessment of progress on changing STEM instruction

for alternatives.” Noted another group, interdisciplinary work can “force us to think differently about instruction. It helps reduce content in intro courses and raises questions about how to do active learning across disciplines.” Finally, collaborative interdisciplinary work was recognized as having an indirect influence on change when it shaped the views of local change agents or encouraged them to take action.

Local Leaders and Internal Change Agents with a Vision

As a lever of change, local leaders acting as change agents were recognized as having high influence across disciplinary groups. “If you don’t have a local change agent to enact the vision and rally support, who will lead?” asked one participant. “All politics are local,” noted another. “This is where the action has to be!” Groups identified some tactics and functions of local leaders that echo those noted in the literature (Kezar, Bertram Gallant & Lester, 2011; Laursen, 2016), including:

- Connect people who are teaching particular courses to communities of practice, and to new ideas;
- Model attitudes and conversations for their colleagues;
- Call for conversations in the department about pedagogy and strategy;
- Connect the department to campus expertise on teaching, learning, and assessment;
- Communicate with senior administrators to understand how to align department goals with institutional interests and what evidence will be persuasive to leaders;
- Mobilize local resources for pilot projects (“Mini-grants of $5000 or less seem to get a lot of effort out of people.”);
- Raise awareness and elevate the status of RBIS studies and practices through local communication mechanisms, such as department meetings, newsletters, on- and off-campus publicity, serving as “a rudder for messaging”;
- Identify opportunities to shift local policies and reward and recognition structures; and
- Curate information about external pressures and opportunities, including workforce issues.

To be effective as “organizational catalysts” (Sturm, 2007), change leaders need to hold informal influence and have access to formal authority. If they are not empowered to enact change, “change is stifled.” Such leadership roles must be elevated and recognized, not seen only as service. Change agents can include instructors, chairs and heads, staff, administrators, alumni, and advisory board members. In some departments, education specialists or DBER scholars might take a leadership role, but it was not always the case that such leadership was part of their job description and expertise. Students could also be change agents; examples include the peer leaders in Peer-Led Team Learning (pltis.org) and Learning Assistants (https://www.
Their work carries RBIS directly to undergraduates, and their observations have helped to improve teaching (Gafney & Varma-Nelson, 2008). Student enthusiasm drove development of an online version of PLTL (J. Smith, et al., 2014). It is important to solicit broad involvement—inventing everyone to the table, not just a select group. Broad involvement is also a way to spread the initiative and develop new leadership, since turnover in faculty or administrators can cause risks to sustainability if the change relies too much on one person’s motivation and input:

We need to invite all faculty to participate and feel like they are part of the effort, to invite them into the identity of student-engaged teaching. We can’t just draw a line around the DBER people and isolate the others.

It is clear we have more to learn about how to support and empower change agents, and about the strategies that can be effective at the local level.

LEVERAGE ON THE STEM EDUCATION COMMUNITY

The groups had high consensus that federal and private funders’ investments in STEM education are very important, “among the strongest” of levers. However, in the discussions it became clear that this final lever operates on a different level than all the others. Funding does not work directly on instructors, nor on academic units, but rather exerts secondary force on the other levers by providing financing, impetus, and legitimacy for activities such as education research, professional development, materials development, educational tools and technologies, and conferences. The National Science Foundation is the major public funder cited in this discussion; in some fields, the Department of Energy, National Institutes of Health, and NASA also invest

in undergraduate STEM education. Private funders mentioned as part of this system include the Howard Hughes Medical Institute and the Research Corporation for Science Advancement, and others could be named. Professional societies were also seen as playing a role, for example, when coordinating funding from industry and private donors.

Funders’ investments have been “crucial for the creation, implementation and legitimization of discipline-based education research, scholarship of teaching and learning, and other research supports for curriculum advancement,” noted one group. Another stated, “Many transformational changes to instruction, and even to institutional culture, have been initiated in response to calls for transforming instruction.” A third concurred: “Funding opportunities have created incentive and opportunity for innovation and resources for some degree of spread.” Funder support for professional development was seen as especially high leverage in raising instructors’ awareness and classroom application of education research findings, and thereby reaching many students.

While financing for entrepreneurial efforts is important, agencies also exert leverage through their programming priorities and designs. For example, program announcements for the National Science Foundation’s ADVANCE program show evolution in the program’s theory of change and communicate evolving expectations and priorities to proposers (Laursen & De Welde, 2019).

Participants suggested some ways that funders might be more influential. One idea was to offer targeted calls for cross-institutional adoption of good ideas, perhaps analogous to the NSF ADVANCE Adaptation and Implementation funding track. Concerns
raised about funder investments as a lever of change included:

- **Sustainability and scale.** The goal to transform STEM education is broad; to have significant impact, change efforts will require long-term engagement well beyond typical funding cycles and funder interest. “We lose focus when we have to chase the next project without seeing the work all the way through.”

- **Reward structures.** Innovators are rewarded for innovating, but it is a different skill set to sustain than to build, and institutional reward structures are less supportive of those who do this work.

- **Inclusion.** “There are still pockets of the community that are not engaged in the funding process.” Two-year colleges and minority-serving institutions were among those mentioned. Some institutions lack the infrastructure and experience that helps to secure and manage grants, and they do not have ready mechanisms to re-allocate faculty time that is ordinarily focused on teaching and advising.

- **Networks and partnerships.** Recent NSF calls have emphasized networks of people and institutions. While these structures are encouraged with good intent to foster collaboration, it can also lead to increased involvement of an in-group, so that “the rich get richer.” Attention and support must be given to the hard work of building inclusive networks, perhaps through mechanisms such as planning grants and conference grants. It may also be important to maintain some support for smaller, entrepreneurial projects, particularly for new investigators or those at institutions with less prior funder support, as done within some NSF programs (Advancing Technological Education, IUSE for Hispanic-Serving Institutions).

Overall, the high level of U.S. funders’ investments in undergraduate STEM education was seen as an essential force behind the progress that has been made to date and that is reflected in the literature reviews and working group assessments.

**OTHER LEVERS FOR CHANGE**

In their discussions, the groups identified additional levers for change. Since these were not discussed by all the groups, they are listed here but not categorized in Table 8.2:

- **Institutional support.** “The chancellor, provost and/or dean articulates explicit and sustained support for this work, backed up by institutional funding for resources and personnel.” Building this support is a critical skill for change leaders.

- **Inclusion of contingent faculty.** “Particularly in developmental and foundational courses, their role is critical and needs to be a focus of attention and resources. We are never going to make progress in instruction and learning in these courses and for these students if contingent faculty continue to be sidelined in terms of instructional support, professional development, pay, and status.” Where non-tenure-track instructors are teaching large introductory courses, some are leaders in adopting RBIS, and others do not have support to learn and adapt these methods. Instructors of online courses should be included in local teaching initiatives as well. Full inclusion here also means advocacy by permanent faculty for equal voice and improved working conditions for colleagues in contingent positions: “It’s a systematic problem that universities are relying on temporary faculty who are not paid enough and
do not have enough job security to be invested in quality teaching.”

**Inclusion of students.** “We can engage students as partners and change the conversation with students to embrace their learning for value as opposed to taking a check-box approach and focusing on grades.... To graduate as competent, self-directed learners, they must know how learning works, and this should be explicitly taught by each teacher.” Students can also be advocates for change, assistants in enacting it, and allies in reducing resistance of their peers.

### PROMISING ARENAS FOR FUTURE WORK

Each of the levers discussed may be part of an overall change strategy—and none of them is effective alone. If we recognize STEM higher education as a complex system of interlocking parts and sub-systems, it is clear that approaches to the uptake problem must be combined strategically. Challenges and opportunities will manifest differently in different parts of the system, and multiple approaches will be needed to exert leverage throughout the system (Laursen, Austin, Soto & Martinez, 2015). The levers for change should address both individual and environmental levels, and they should include a mix of approaches that encourage the use of well-developed practices and involve people in analyzing data, reflecting on their situation, and making decisions collectively (Beach, Henderson & Finkelstein, 2012). “Approaching the same cause from multiple angles is critical to success!”

From the working groups’ assessments of their own fields, we identified and distilled eight “promising arenas” for future work to expand and deepen the impact of research-based instruction. Meeting participants chose one of these arenas to discuss with peers and met in discipline-mixed breakout groups to discuss the next steps in each arena. Their discussions were richly cross-fertilized by the previous exchange of ideas, and their ideas for “next steps” in each arena offer a rich set of possibilities for individual leaders and disciplines, institutions, professional organizations and funders to consider. As before, quotations are drawn from the working documents and verbal reports as recorded in meeting transcripts and are not attributed to particular speakers. The perspectives portrayed here are based on qualitative analysis of the breakout groups’ notes and presentations; relevant literature has been cited to offer additional validation for these ideas as promising arenas for future work.

#### Addressing Institutional Differences in Roles and Opportunities

It is important to attend to institutional and disciplinary context. For example, recognize differences in what motivates instructors at PUIs or research universities, to understand how to apply leverage differently in varied settings to nurture and validate champions for research-based instruction, and to deploy distinct institutional reward structures to engage instructors widely. Regional comprehensive institutions, PUIs, minority-serving institutions, liberal arts colleges, and two-year colleges are important players with distinct cultures and missions who educate large numbers of diverse students, but they are often underrepresented in DBER work and institutional change efforts (e.g., NASEM, 2019). Indeed, growing emphasis by funders on generating research as part of educational development projects may further favor research-oriented institutions. Ideas for next steps include:

- Measuring and understanding patterns around institutional type in the uptake of
RBIS and the levers for change available in different institutions to enhance equitable, research-based instruction;

▲ Fostering partnerships and professional development to co-engage instructors from 2YCs and other types of institutions with education research (see, e.g., Schinske, et al., 2017); and

▲ Recognizing the first two years of STEM curricula as a unifying space for educators in 2Y and 4Y institutions and convening groups to work on this.

**Developing Departments as Loci of Change**

Developing formal and informal leadership is key to working in departments, by activating department chairs as collaborators and by developing effective informal leaders and advocates within departments. Recent research is also beginning to identify conceptual and practical strategies that can be used to guide department-based change (see e.g. Andrews, Conaway, Zhao & Dolan, 2016; Apkarian, Bowers, O’Sullivan & Rasmussen, 2018; Henderson, et al., 2018; Laursen, 2016; Quardokus Fisher & Henderson, 2018; W. Smith, Webb, Bowers & Voigt, 2017). Possible next steps include:

▲ Working with mid-level institutional leaders, such as department chairs, heads and deans, to increase their sense of urgency and agency around inclusive and effective STEM instruction and to help them recognize the levers for change within their grasp;

▲ Building leadership capacity of informal opinion leaders who can serve as change agents, drawing on models such as the PULSE community (www.pulse-community.org); and

▲ Developing practical resources and toolkits for change leaders on what they can do and how to do it (for an example from a different domain, see Laursen & Austin, 2014).

**Creating Linkages Among Education Researchers and Classroom Practitioners**

“We can’t expect all instructors to be connected with DBER and we can’t expect all DBER researchers to be connected and disseminating the results, but we need people and processes from both communities that work together, rather than expecting every community to do this work.” Possible next steps include:

▲ Developing research coordination networks in and across STEM fields, like those now supported by NSF in biology (Diaz Eaton, et al., 2016), with a particular eye toward fostering research-practice collaborations that engage diverse people and institutions;

▲ Target professional development and leadership development to those who hold essential bridging roles, such as course coordinators, people who train graduate teaching assistants, and faculty developers;

▲ Raising the professional status of activities that communicate education research to practitioners, through cross-disciplinary connections among DBER communities or through mechanisms such as CAREER awards and broader impacts in NSF projects; and
Generating more ‘scholarship of translation’ writing and places to publish it.

**Making Inclusion a Driver, not a Follower, of Instructional Change**

To achieve this, we must redefine effective teaching as inclusive teaching. While many studies show that active learning can benefit students on the whole, students with marginalized identities may be at heightened risk and may experience the classroom quite differently (e.g., Cooper & Brownell, 2016; Linley, Renn & Woodford, 2018; Van Dusen & Nissen, 2017). Students, faculty, professional societies, foundations, funders, industry, all have a stake in this process, so “how can we bring together all those players and create urgency for each stakeholder?” Ideas for next steps at classroom and institutional levels include:

- Learning from research on students, and from our own students, what shapes the learning environment into one where they feel they belong in all their intersectional identities, and how courses need to be designed and delivered differently;

- Helping instructors build a better understanding of what classrooms feel like for students, gather data about inequities they may not recognize in their own courses and attend to their relational interactions (see Battey, 2013; Battey & Leyva, 2018);

- Developing and using data that go beyond measures of content knowledge mastery, to measure not just equality in outcomes but equity in experiences;

- Marshalling and comparing institutional data to reveal and explore inequities in student outcomes and pathways; and

- Deploying tools such as departmental plans for diversity, equity and inclusion, to increase accountability and reward successful programs.

**Shifting Instructional Mindsets from Deficit to Asset Models**

This fundamental shift in thinking, from an instructor mindset that is content- or instructor-centered to one that is student-centered and taking responsibility for student success, must accompany the use of RBIS. Recent studies show that instructor beliefs about students’ abilities are closely tied to their use of RBIS and to the quality and equity of their students’ achievement and experiences (Aragón, Eddy & Graham, 2018; Canning, Muenks, Green & Murphy, 2019). An asset-based mentality recognizes the talent, skills, capacities and strengths of individuals and communities that can be mobilized for action (Mein, 2018). Potential next steps include:

- Developing instructors’ awareness and capacities to align their practices with an asset-based approach such as assessments that enable students to display their assets; opportunities to make decisions, experience failure, reflect and learn from it; and grading strategies that de-emphasize curved exams and move toward competency-based assessment;

- Providing implicit bias and other trainings for instructors to consider how bias may play out in the classroom and strengthen skills in catching and interrupting their own and others’ biases and disrupting microaggressions (Battey, 2013; Battey & Leyva, 2018; Moss-Racusin, et al., 2016);

- Examining institutional data, websites, environments and processes for deficit messages and unequal outcomes (see, e.g., Matz, et al., 2017); and
Tracking data to identify how an assets-based approach can benefit the institution by retaining students and producing happy alumni. For example, evidence about lower costs and efficient spending may be helpful in speaking with senior institutional leaders and constituencies of public institutions.

**Addressing Gendered Participation in Teaching and Learning Work**

Women are routinely represented in STEM education and equity work beyond their numbers in the profession, and they disproportionately hold positions, such as non-tenure-track instructor, that are lower in status but essential to departments' teaching work. Likewise, the intellectual work they do often has low visibility and status: mentoring students and colleagues, leading departmental planning and assessment work, and advising or supporting student-centered initiatives or groups. This work is often counted as service, outreach, or not at all, rather than valued as scholarly and creative. Advocacy for change must recognize and actively counter the devaluing of this work. Next steps may include:

▲ Raising awareness and taking inventory of hidden and emotional work and how it manifests in STEM environments and in gendered ways;

▲ Developing strategies for meaningful partnering of instructional change efforts with educational researchers and scholarly teachers that engage their expertise, but do not reduce their efforts to outreach or service; and

▲ Including and empowering those in lower-status positions to contribute to departmental decision-making and to develop as leaders in STEM instruction, including course coordinators, undergraduate advisors, laboratory instructors, and other academic professionals.

**Expanding and Deepening the Role of Professional Societies**

One challenge for societies is the predominance of research in their agendas, which results in marginalizing education within the society. Potential next steps within societies include:

▲ Enhancing support for professional development, including workshops on research-based instruction for new (and experienced) faculty, and leadership development of change leaders;

▲ Increasing the visibility and prestige of society awards for education research and leadership;

▲ Gathering and evaluating data on instruction and student outcomes, and using it to craft a “data-driven agenda” to guide programming and resource allocation for improving instruction and equity; and

▲ Being mindful and proactive about messages and images about instruction that have symbolic impact in the disciplinary community, such as attending to the images of classrooms and teaching that are used on websites and the covers of reports.

**Finding and Creating Tools to Enhance Conversation Across Disciplines**

This approach requires building STEM education leaders’ capacities to communicate with broader audiences, so that important messages about effective instructional practices reach a wide range of audiences. “Leadership development can’t wait until the moment you become department chair!” Equally important, but operating at a different level, is connecting people...
locally across departments and offices on campuses, which may include teaching and learning centers, academic support, student affairs, institutional research, and others who take interest in student success. “There are limits to what a physics department at an institution can do unless other STEM departments are also involved.”

This raises the profile of effective teaching efforts, supports individuals, and helps to identify synergies and reduce redundancies in campus efforts. Possible next steps include:

▲ Building toolkits and professional development opportunities to help instructors strengthen their communication and advocacy skills. Examples of work that may provide models include the COACh workshops (https://coach.uoregon.edu/), workshops for senior women change agents (e.g. https://www.aps.org/programs/women/workshops/skills/), and the Alan Alda Center for Science and Medicine Communication (https://www.aldacenter.org/);

▲ Connecting education researchers to investigate lexical differences between disciplines around the topic of instructional practices. Also develop resources to help resolve or unify these differences for non-specialists; and

▲ Fostering a culture of cross-discipline conversation through local “support groups for research-based teachers.” This can include using mechanisms such as journal clubs, mailing lists, events, workshops, peer observations, campus teaching academies, and other forums.

CONCLUSION

Taken together, the findings show a sense of accomplishment in progress made and a sense of optimism that further progress is possible. These are coupled with high awareness that the next steps for STEM instruction involve collaborative, multi-level approaches, moving beyond simple propagation strategies of the past that assume methods can be developed and handed off (Henderson, et al., 2015). They reflect an emerging organizational perspective that acknowledges the need for theories of change that address the systems and structures where STEM instructors work, as well as the instructors themselves (Austin, 2011; Beach, Henderson & Finkelstein, 2012; Elrod & Kezar, 2016; Seymour, 2002). Importantly, they recognize that purposeful work is needed to enable these higher-order approaches: building leadership capacity and cross-cutting communication skills in disciplines and institutions; planning for scalability and sustainability both in scholarship and in sharing existing wisdom; attending mindfully to messages and symbols as well as the content of initiatives to enhance STEM instruction. Such organizational perspectives will be essential to drive leaders’ learning and foster the new types of collaborations needed to infuse equity concepts and classroom strategies into professional development and campus conversations about STEM instruction.

Ultimately, “change is less about the thing being changed ... and more about changing beliefs about teaching and learning” (Giersch & McMartin, 2014, p. 8). At heart, all these levers can be viewed as ways to reach instructors and influence what they believe about teaching and learning. That is not to say that instructor desire alone makes the difference, but rather, as individuals'
understandings of learning and commitments to student-centered teaching take root, it also becomes more possible to mobilize leaders, to garner support to remove the situational factors that constrain use of RBIS, and to enhance the cultures and structures that support it (Henderson & Dancy, 2007). Positive feedback loops develop as shifts in cultures and structures encourage and support more individuals to make changes that will benefit their students.

Many of these levers for change have a long history: Today’s projects and leaders are not the first to take action on these challenging problems. But in the past decade, levers at different levels and from different directions have come to work less in isolation and more in synergy. We can now recognize their combined effect as a significant, encouraging and irreversible shift in instructors’ use of these research-based approaches to teaching. What these discussions of the state of STEM instruction today and tomorrow also reveal is a sense of optimism, collegiality, and support more individuals to make changes that will benefit their students.

REFERENCES


Aragón, O. R., Eddy, S. L., & Graham, M. J. (2018). Faculty beliefs about intelligence are related to the adoption of active-learning practices. CBE—Life Sciences Education, 17(3), ar47.


Canning, E. A., Muenks, K., Green, D. J., & Murphy, M. C. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. Science Advances, 5(2), eaau4734.


President’s Council of Advisors on Science and Technology (PCAST) (2012). Report to the President. Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. Washington, DC: Executive Office of the President, President’s Council of Advisors on Science and Technology.


184 LEVERS FOR CHANGE: An assessment of progress on changing STEM instruction


May 7-8, 2018
HHMI Headquarters
Chevy Chase, MD

MEETING ORGANIZERS/FACILITATORS
Sandra Laursen, Project Lead
University of Colorado Boulder
Ann Austin
Michigan State University
Beth Ruedi
AAAS

CHEMISTRY
Renee Cole
University of Iowa
Laurie Langdon
University of Colorado Boulder
Susan Rowland
The University of Queensland
Brisbane, Australia
Kristin Plessel
University of Wisconsin Colleges
Susan Shadle, Convener
Boise State University
Marilyne Stains, Reviewer
University of Nebraska-Lincoln
Pratibha Varma-Nelson
IUPUI
Leyte Winfield
Spelman College

EARTH SCIENCES
Leilani Arthurs
University of Colorado Boulder
Eric Baer
Highline College
Heather Macdonald
College of William & Mary

David McConnell, Reviewer
North Carolina State University
Eric Riggs, Convener
Texas A&M University
Karen Viskupic
Boise State University
Emily Ward
Rocky Mountain College

ENGINEERING
Maura Borrego, Reviewer
University of Texas at Austin
Cindy Finelli, Reviewer
University of Michigan
Ann Q. Gates
The University of Texas at El Paso
Jill Nelson
George Mason University
Olga Pierrakos, Convener
Wake Forest University
Ron Ulseth
Iron Range Engineering
Julia Williams
Rose-Hulman Institute of Technology

LIFE SCIENCES
Tessa Andrews, Reviewer
University of Georgia
A. Malcolm Campbell, Convener
Davidson College
Alix Fink
Longwood University
J.K. Haynes
Morehouse College
Ann J. Murkowski
North Seattle College
Jeff Schinske
Foothill College
Michelle Smith
University of Maine

Kimberly Tanner
San Francisco State University

**MATHEMATICAL SCIENCES**

Rick Cleary
Babson College

Ann Edwards
WestEd/Carnegie Math Pathways

Estrella Johnson, Reviewer
Virginia Tech

Shonda Kuiper
Grinnell College

Rachel Levy, Convener
Harvey Mudd College

Natasha Speer
University of Maine

David Webb
University of Colorado Boulder

**PHYSICS/ASTRONOMY**

Kathleen Foote, Reviewer
University of British Columbia
Vancouver, Canada

Charles Henderson
Western Michigan University

Michael Jackson
Millersville University of Pennsylvania

Julia Kregenow
Penn State University

Timothy McKay, Convener
University of Michigan

Craig Ogilvie
Iowa State University

Scott Schultz
Delta College

**MULTIPLE FIELDS/OTHER**

Myles Boylan
Retired, National Science Foundation

Shirley Malcom
AAAS

Emily Miller
Association of American Universities

Kacy Redd
Association of Public and Land-grant Universities

**NSF OBSERVERS**

Ellen Carpenter
Division of Undergraduate Education
National Science Foundation

Connie Della-Plana
Division of Undergraduate Education
National Science Foundation

Muriel Poston
Division of Biological Infrastructure
National Science Foundation

**HHMI STAFF – COORDINATORS**

Sarah Simmons
Assistant Director Science Education
HHMI

Josephine Bynum
Program Assistant Science Education
HHMI

Melvina Lewis
Program Coordinator Science Education
HHMI

**HHMI STAFF – OBSERVERS**

David Asai
Senior Director for Science Education
HHMI

Susan Musante
Program Officer Science Education
HHMI

Andrew Quon
Senior Program Officer Science Education
HHMI

Caitlin Schrein
Science Education
HHMI

Patricia Soochan
Program Officer Science Education
HHMI
Levers for Change: Reviewers and Conveners

REVIEWERS

Tessa Andrews, Life Sciences
University of Georgia

Maura Borrego, Engineering & Comp. Sciences
University of Texas at Austin

Cindy Finelli, Engineering & Comp. Sciences
University of Michigan

Kathleen Foote, Physics & Astronomy
University of British Columbia
Vancouver, Canada

Estrella Johnson, Mathematics & Statistics
Virginia Tech

David McConnell, Geosciences
North Carolina State University

Marilyn Stains, Chemistry & Biochemistry
University of Nebraska-Lincoln

CONVENERS

A. Malcolm Campbell, Life Sciences
Davidson College

Rachel Levy, Mathematics & Statistics
Harvey Mudd College

Timothy McKay, Physics & Astronomy
University of Michigan

Olga Pierrakos, Engineering & Comp. Sciences
Wake Forest University

Eric Riggs, Geosciences
Texas A&M University

Susan Shadle, Chemistry & Biochemistry
Boise State University
Levers for Change
Working Meeting: Agenda

May 7-8, 2018
HHMI Headquarters
Chevy Chase, MD

MEETING GOALS
To capture a snapshot of the current state of research-based reform in undergraduate STEM instruction within six clusters of STEM disciplines

- biological sciences
- chemistry and biochemistry
- engineering and computer science
- geosciences
- mathematical sciences
- physics and astronomy

To identify key levers of change that are seen to have been effective in reaching this state, and to identify additional levers—less-tapped or untapped—that may be useful in the next decade

To convene a group of leaders in research and practice to learn from, inspire and connect with each other

DRIVING QUESTIONS
These questions guide our agenda:

- What is the current state of research-based reform in undergraduate instruction within six clusters of STEM disciplines?
- How did each arrive there? What factors or events have been important in reaching this state? And how are these similar or different by discipline?
- What provides evidence for these trajectories of change, and why?
- What can be learned from this evidence about how to expand and deepen the impact of these changes in the next decade?

WORKING PREMISES

- There is sufficient evidence from education research in and across the disciplines to indicate that active learning experiences are good for students in supporting their learning, attitudes, sense of belonging and persistence in STEM. (We know that ongoing studies will further detail these benefits and how they vary for different students and settings.)
- We take a broad view of what strategies count as active learning. (We acknowledge that there is more to learn in comparing and contrasting strategies, their affordances and limitations.)
- We focus on classroom instruction provided by college instructors in STEM disciplines. (We recognize that co-curricular and extracurricular experiences also matter for students.)
- Instructors’ uptake of active learning strategies is a critical step in ensuring that all students experience the benefits of active learning in their STEM courses. Their skillful and sustained use of these strategies is a necessary next step.
- We use the term ‘evidence’ inclusively, considering scholarly evidence from research and evaluation but also experience and observations from the field.
MONDAY, MAY 7

SETTING THE STAGE: WHO ARE WE AND WHAT ARE WE DOING HERE?

2:00 PM, Day 1 Opener, in D124-125. Welcome, scene-setting, and introductions. Sandra Laursen, Beth Ruedi & hosts.

DRIVING QUESTION 1:
What is the current state of research-based reform in undergraduate instruction within each STEM discipline?


The disciplinary reviewers will share their answers to Driving Question 1 (DQ1), based on the scholarly literature in their field. We’ll take clarifying questions after each presentation, then have time for open questions and discussion.

**Biological sciences**: Tessa Andrews, U. Georgia

**Chemistry & biochemistry**: Marilyne Stains, U. Nebraska Lincoln

**Engineering/CS**: Cindy Finelli, U. Michigan, & Maura Borrego, U. Texas Austin

**Geosciences**: David McConnell, North Carolina State

**Mathematical sciences**: Estrella Johnson, Virginia Tech

**Physics/astro**: Katie Foote, U. British Columbia

Charge to the working groups.

4:15 PM Break—Refreshments available in the Great Hall. Move to breakout spaces in Disciplinary Groups with conveners.

**Biological sciences**: Malcolm Campbell, Davidson—D115

**Chemistry & biochemistry**: Susan Shadle, Boise State—D116

**Share D124-125, sites as viewed from the panelists’ position:**

**Engineering/CS**: Olga Pierrakos, Wake Forest—rear, stage L

**Geosciences**: Eric Riggs, Texas A&M—rear, stage R

**Mathematical sciences**: Rachel Levy, Harvey Mudd—front, stage L

**Physics/astro**: Timothy McKay, U. Michigan—front, stage R

4:30 PM Work Session 2, in breakout spaces. Disciplinary Groups characterize the state of uptake of active learning instruction in their own cluster of disciplines

**Group task**: Take a few minutes for introductions, then discuss Driving Question 1 (DQ1) and your takeaways from the white papers. Focus on your own field for now.

**For discussion**: What is your assessment of the state of uptake in your field so far? Where has the field made progress so far? Where has it lagged? What is your evidence? If your responses vary by student audience, course, instructor or institution type, etc., please capture such nuances in the notes.

To what extent is the group in consensus, or not, about these views?

What evidence can group members add? – examples, counter-examples, important projects or initiatives.

In addition to group notes from the breakout sessions, please capture individuals’ additions in the file “other ideas” in your Disciplinary Group’s folder. Include details, links or citations and explain what it says or why you think it is important. Include your name so we can follow up if needed. You can upload files here too.

5:30 PM Groups take stock and review the questions that you will answer on a poster
(see Work Session 3). You will have work time to create your poster after dinner.

5:45 PM Reception in the Great Hall

6:15 PM Dinner in the Dining Room
Please sit at the table matching the colored dot on your nametag. This is your Jigsaw Group, mixed by disciplines. You will meet in Jigsaw Groups again on Tuesday. You will be responsible to bring ideas from your Disciplinary Group to share in Jigsaw discussions, and to bring ideas back from the Jigsaw to your home group.

As you begin dinner, go around the table for introductions – name, institution, role. Share one ‘aha’ or insight that your disciplinary group discussed.

7:30 PM, Work Session 3. Disciplinary Groups make posters in breakout spaces.

**Group task:** Identify key points of your discussion and prepare a poster using 2 or 3 big sticky notes to answer these three questions:

1 How do you characterize the current state of uptake in your field? Choose a metaphor, a labeled graphic, or 4-6 adjectives to describe this state.

2 What has moved our field forward on uptake of active learning instruction? And why we think so.

3 What we wonder about: Ideas, hypotheses, conjectures, questions about the uptake of active learning instruction

Place your posters on the big poster boards in the Atrium – one disciplinary group on each side of the board.

8:15 PM, Gallery Walk, in the Atrium.
Compare and contrast posters from the disciplines.

**Individual task:** Pick up small sticky notes and review 3 posters besides your own. Contribute a comment on a sticky note to each poster you view. What ideas are intriguing or surprising, similar or different from your own field? What insights are offered, what questions are prompted, by the set of posters as a set?

8:45 PM Return to your group’s poster and review the sticky notes together. Do a quick qualitative analysis to cluster the comments and identify key themes or insights.

9:00 PM Reporters share out themes from the comments on their poster, one minute each.

9:06 PM Conveners meet for a quick debrief.

9:06-10:30 PM Social time in the Pilot Lounge.

~ SLEEP ~

**TUESDAY, MAY 8**

7:30 AM Checkout and Breakfast in the Dining Room.
Please return your room key to the Conference Desk. Luggage may be stored in the Conference Center.

**DRIVING QUESTION 2:**
How did we get here? What factors or events have been important in reaching this state of affairs?

8:30 AM, Plenary session in D124-125:
Panel on cross-STEM perspectives.

A panel of observers with experience of instructional reform across STEM fields will offer observations on Driving Question (DQ) 2: Myles Boylan, NSF (retired); Emily Miller, AAU; Sarah Simmons, HHMI. Scene setting for day 2 and charge to working groups; move to breakout spaces.
9:10 AM, Work Session 4, in breakout spaces. Disciplinary Groups go deeper in discussing levers and barriers to change.

**Group task:** This session focuses on Driving Question 2 to formalize and organize ideas emerging from Monday’s sessions. We will ask you to consider some specific levers of change and what actual or potential influence they have had in the past and could have in the future, in your field. We will also ask you to consider what has held your field back, and why.

**For discussion:** What levers and barriers have been important in your discipline, and why? Which levers active in other fields have not been important in your field, and why?

In responding, please think systemically about how structures and cultures in your field affect instructor uptake of active learning, positively or negatively—e.g., how teaching and learning is organized; how your field is professionally organized; your discipline’s professional norms and ways of knowing. What affordances or constraints do these structural and cultural features offer that affect instructor uptake of active learning instruction? For example: Disciplines often organize courses differently—in large ‘lecture’ sections or many small ones; with lab work, field trips, recitations, or computer labs. How do these structures affect the uptake of active learning by instructors and departments in your field?

10:00 AM Groups stop and prepare report out: one lever that looks most promising as a next step in your field.

10:10 AM Reconvene in D124-125 and groups report out: 1 slide, 2 minutes each.

10:25 AM, Sandra and Ann detail the next session—considering promising arenas for next steps, using ideas generated in the working groups and posters. Sign up for a “promising arena” that interests you. Write your name on a small sticky note and add it to the poster page for the topic that you’d most like to talk about. If you’re willing to be a team captain, please include that on your sticky.

10:45 AM, Coffee break in the Great Hall.

11:15 AM, Work Session 5, in breakout spaces. Meet with promising arena/interest groups.

12:15 PM, Groups stop and prepare to report out (one slide each, afternoon plenary).

12:30 PM, Lunch in the Dining Room

**DRIVING QUESTION 3:**

**What’s next? Opportunities and challenges for the future. What have we learned that can help us expand and deepen the impact of these changes in the next decade?**

1:15 PM, Report out, working groups present their report on their discussion around a promising arena. As you listen to each group’s report, write your response to their ideas on a small sticky note. Collect them at your table and add to the poster page for each topic.

2:00 PM, Plenary discussion, in D124-125.

Where do we go from here? Moving forward.

General discussion on ideas for impact. How do we amplify important messages from this meeting and the report?

2:50 PM Closing comments—thanks and next steps for this project

3 PM, Adjourn
For more information about Levers for Change, please see
https://www.aaas.org/resources/levers-change-assessment-progress-
changing-stem-instruction

This material is based upon work supported by the National Science
Foundation under Grant No. DUE-1141449

Special thanks to the Howard Hughes Medical Institute for providing
their conference facilities and on-site meeting support, which allowed
us to include more participants and their valuable perspectives, greatly
enhancing this report.