A Faculty Professional Development Model That Improves Student Learning, Encourages Active-Learning Instructional Practices, and Works for Faculty at Multiple Institutions

Karen N. Pelletreau,[†‡] Jennifer K. Knight,[§] Paula P. Lemons,[‖] Jill S. McCourt,[‖¶] John E. Merrill,[#] Ross H. Nehm,[@] Luanna B. Prevost,[**] Mark Urban-Lurain,[††] and Michelle K. Smith[†*]

[†]School of Biology and Ecology, University of Maine, Orono, ME 04469; [§]Department of Molecular, Cellular, and Developmental Biology, University of Colorado, Boulder, CO 80309; [‖]Department of Biochemistry and Molecular Biology, University of Georgia, Athens, GA 30602; [#]Department of Microbiology and Molecular Genetics and [††]CREATE for STEM Institute, Michigan State University, East Lansing, MI 48824; [@]Department of Ecology and Evolution, Stony Brook University (SUNY), Stony Brook, NY 11794; **Department of Integrative Biology University of South Florida, Tampa, FL 33620

ABSTRACT

Helping faculty develop high-quality instruction that positively affects student learning can be complicated by time limitations, a lack of resources, and inexperience using student data to make iterative improvements. We describe a community of 16 faculty from five institutions who overcame these challenges and collaboratively designed, taught, iteratively revised, and published an instructional unit about the potential effect of mutations on DNA replication, transcription, and translation. The unit was taught to more than 2000 students in 18 courses, and student performance improved from preassessment to postassessment in every classroom. This increase occurred even though faculty varied in their instructional practices when they were teaching identical materials. We present information on how this faculty group was organized and facilitated, how members used student data to positively affect learning, and how they increased their use of active-learning instructional practices in the classroom as a result of participation. We also interviewed faculty to learn more about the most useful components of the process. We suggest that this professional development model can be used for geographically separated faculty who are interested in working together on a known conceptual difficulty to improve student learning and explore active-learning instructional practices.

INTRODUCTION

Many faculty are aware that their students hold inaccurate ideas about science concepts, and they would like to address these conceptual difficulties in their teaching using active-learning teaching strategies that have been shown to improve student performance and engagement in science, technology, engineering, and mathematics (STEM) classrooms (Freeman *et al.*, 2014). However, lack of time, incentives, motivation, and professional development opportunities are impediments to creating new instructional materials (Silverthorn *et al.*, 2006; Wieman *et al.*, 2010; Anderson *et al.*, 2011; Ebert-May *et al.*, 2011, 2015). Furthermore, simply providing instructional materials to faculty in the absence of a community or guidance does not necessarily result in an alteration in faculty practices (Sharp and McLaughlin, 1997; Penberthy and Millar, 2002; Silverthorn *et al.*, 2006; Henderson *et al.*, 2011), so additional work

Mary Lee Ledbetter, *Monitoring Editor*

Submitted December 11, 2017; Revised March 12, 2018; Accepted March 13, 2018

CBE Life Sci Educ June 1, 2018 17:es5

DOI:10.1187/cbe.17-12-0260

Present addresses: [†]Center for Innovation in Teaching and Learning, 102 Fernald Hall, University of Maine, Orono, ME 04469; [¶]Chemistry Department, University of Puget Sound, Tacoma, WA 98416.

*Address correspondence to: Michelle K. Smith (michelle.k.smith@maine.edu).

is needed to understand the professional development supports that promote instructional change.

We worked to overcome many of these challenges by involving faculty, who voluntarily participated in a professional development opportunity, in the iterative design of an instructional unit that uses active-learning pedagogy focused on conceptual difficulties related to the central dogma of biology. Specifically, we explored the efficacy of using student learning data to motivate faculty change, a facilitator as a way to protect faculty time and organize discussions around student learning data, and a collaborative publication to incentivize faculty involvement. We asked, 1) How can we minimize time investment for faculty yet engage faculty to develop an instructional unit with an active-learning approach that positively affects student understanding? 2) How does involvement in this process influence faculty willingness to try new instructional practices? To answer these questions, we used a design-based research approach (reviewed in Anderson and Shattuck, 2012) in which participating faculty designed and made iterative changes to an instructional unit based on student assessment data.

We designed our efforts based on professional development models that have been reported to increase faculty use of active learning. One model we used was the Summer Institutes on Scientific Teaching (SI; www.summerinstitutes.org), in which participants learn about scientific teaching, active learning, assessment, and inclusive teaching in a weeklong immersive professional development (Pfund *et al.*, 2009). The participants also develop a "Teachable Tidbit"—an instructional unit to be used at their home institutions (Wood and Handelsman, 2004). SI faculty self-report an increased use of active learning in their classrooms (Pfund *et al.*, 2009). Additionally, many of the SI faculty disseminate scholarship that arises from their SI experience; 25% of faculty who participated in the first 5 years of the program (~50 faculty) published manuscripts about their instructional units (e.g., Hoskinson *et al.*, 2014; Sestero *et al.*, 2014; Emtage *et al.*, 2016; Freeman *et al.*, 2017) and/or teaching efforts (Pfund *et al.*, 2009).

We also drew inspiration from faculty learning communities (FLCs). FLCs are networks of eight to 16 faculty members who work together over several months (Cox, 2004, 2016). Several permutations of FLCs exist, with variations in size, frequency of meeting, and goals; however, the long-term engagement by faculty is the key element (Thompson *et al.*, 2015). Because change in faculty instructional beliefs and practices can occur slowly (Derting *et al.*, 2016), FLCs allow repeated practice and reflection and provide the opportunity to discuss and implement change as a part of a group, rather than in a vacuum (Ebert-May *et al.*, 2011; Henderson *et al.*, 2011). One large-scale study of FLC participants found that 79% self-reported at least a moderate, and in some cases a substantial, improvement in student learning based on their FLC participation (Beach and Cox, 2009). The long-term nature of FLCs can also support participation in the scholarship of teaching and learning through faculty dissemination of their experiences in presentations or publications (Richlin and Cox, 2004).

In an initial meeting of faculty associated with this project, the faculty united around the collective discovery that their students were struggling with concepts related to the central dogma of biology (student responses described in detail in Table 1 later in this paper). Through the use of constructed-response questions about the influence of a stop-codon mutation on DNA replication, transcription, and translation (Prevost *et al.*, 2016), these faculty learned that many of their students had a combination of correct and incorrect ideas, referred to as "mixed mental models" (Opfer *et al.*, 2012; Prevost *et al.*, 2016). Although these conceptual difficulties had been previously identified, very few instructional resources existed to help instructors address them in the classroom (Smith *et al.*, 2008; Smith and Knight, 2012; Wright *et al.*, 2014).

To address students' mixed mental models, 16 faculty from five different institutions collaboratively developed an instructional unit similar to the Teachable Tidbits produced at the SIs (Wood and Handelsman, 2004) and taught the instructional unit over several semesters. The community of faculty met virtually over several semesters to facilitate the sharing of data and teaching experiences, iteratively make data-driven changes to the instructional unit based on student learning results, and disseminate the final product (Pelletreau *et al.*, 2016). In addition, at each institution, the faculty members met regularly in groups, also inspired by the FLC model, that focused on helping faculty implement formative assessment questions in their classes (McCourt *et al.*, 2017). Here, we share a combination of student learning, classroom observation, and faculty interview data to demonstrate that this professional development model connected faculty who were working across multiple institutions, helped them use evidence and data from their classrooms to iteratively design a new instructional unit that positively affected student learning, and encouraged them to use more active-learning instructional techniques in class.

For the student data, approval for this study was obtained from the following institutional review boards: University of Georgia: Study 00000256; University of Maine: Study 2012-12-14; Michigan State University: Study x10-577; University of Colorado Boulder: Study 0610.10; Stony Brook University: Study 504271-3. For the faculty data, the University of Georgia IRB board approved this study under exempt status, Study 00000257.

FACULTY PARTICIPANTS

The faculty who participated in this project all taught large-enrollment biology courses at research-intensive PhD-granting universities and were members of the Automated Analysis of Constructed Response (AACR) project. The AACR project focuses on developing computer resources for automated scoring of constructed-response short-answer assessment items intended for formative assessment in large-enrollment undergraduate STEM courses (Haudek *et al.*, 2011, 2015; Moharreri *et al.*, 2014; Prevost *et al.*, 2016; https://msu.edu/~aacr).

As part of the AACR project, faculty engaged in local professional development groups inspired by FLCs (Cox, 2004, 2016), in which they met three times per semester with a discipline-based education research (DBER) faculty member and colleagues (Figure 1) to discuss constructed-response questions they were asking, student responses to the questions, and changes they would like to make to instructional practices (McCourt *et al.*, 2017). All faculty were originally asked to be part of the local AACR groups because they teach large-enrollment biology courses at their respective institutions. When interviewed about why they joined the AACR project, several faculty reported valuing the opportunity to talk with colleagues about

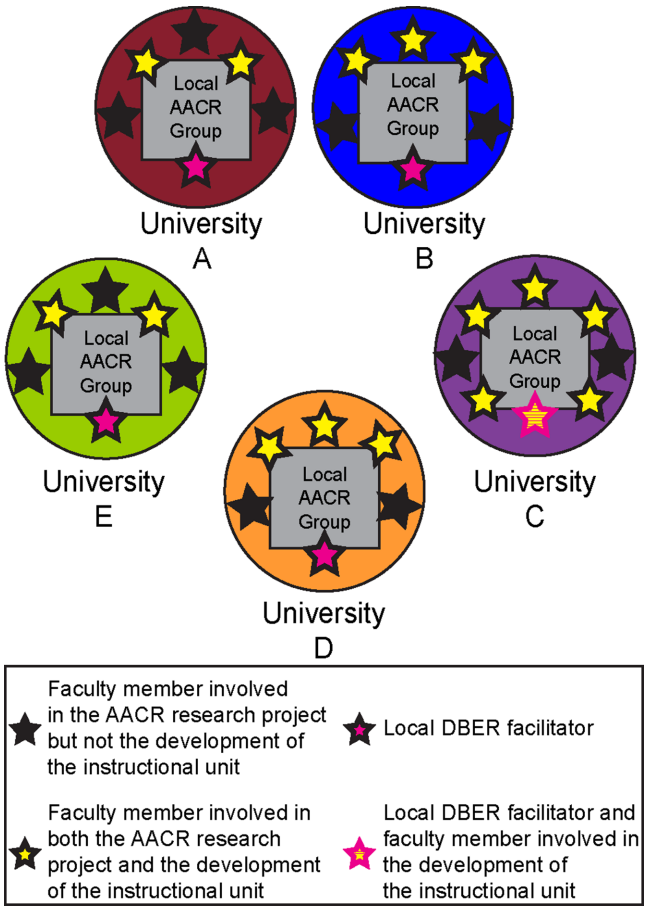

FIGURE 1. The faculty professional development group that designed the instructional unit included 16 faculty (yellow stars) who also participated in local AACR groups (all stars) at five different universities. Each *local* AACR group was facilitated by a DBER faculty member (pink stars). One DBER facilitator also participated in the development and teaching of the instructional development unit (yellow and pink striped star). The number of faculty involved in the local AACR groups but not in the instructional unit development (black stars) varied between institutions. The AACR community included one additional university, but none of their members participated in the effort described here.

The following DNA sequence occurs near the middle of the coding region of a gene.

DNA 5′ A A T G A A T G G* G A G C C T G A A G G A 3′

There is a G to A base change at the position marked with an asterisk. Consequently, a codon normally encoding an amino acid becomes a stop codon.

1. **How will this alteration influence DNA replication?**
2. **How will this alteration influence transcription?**
3. **How will this alteration influence translation?**

FIGURE 2. The AACR stop-codon assessment questions asked before and after different versions of the instructional unit.

teaching and an interest in the education research projects led by the local DBER faculty member (McCourt *et al.*, 2017).

A subset of the AACR faculty (Figure 1, yellow stars) asked their students open-response questions about the effects of a mutation that results in a premature stop codon on DNA replication, transcription, and translation (Prevost *et al.*, 2016; Figure 2). Student written responses are categorized as correct, irrelevant/unclear, or incorrect by the AACR algorithms that provide a more detailed understanding of student thinking about these complex processes. Examples of student responses for each of these categories are provided in Table 1. Data from multiple institutions show that ~45% of student answers are classified as incorrect or irrelevant/unclear, even *after* instruction on the central dogma of biology (Prevost *et al.*, 2016).

The student answers inspired faculty in this project to work collaboratively to develop an instructional unit to improve student understanding of mutations and the central dogma of biology. Additional information about the faculty who participated in the project, including their years of teaching experience and self-selected role in the group, is shown in Supplemental Table 1. To protect the identity of the participants, all faculty have been given pseudonyms.

STRUCTURE OF THE FACULTY PROFESSIONAL DEVELOPMENT

The faculty collaboration started with an in-person meeting, which was funded by a National Science Foundation (NSF) WIDER grant, followed by several virtual meetings. At the in-person meeting, faculty self-selected into groups based on the AACR questions they were asking in their classes (e.g., genetics, evolution, photosynthesis). These groups included faculty from multiple institutions. The faculty who asked the genetics stop-codon questions (Figure 2) met together to talk about student responses to these questions and brainstorm ideas for in-class activities that could help students with common conceptual difficulties identified using the questions. One idea included a case study in which students would explore nucleotide differences in two different individuals and answer questions about how the introduction of a premature stop codon would impact various stages of the central dogma of biology. Several faculty members were using clickers or were interested in trying clickers, so they also began to write multiple-choice questions that could be used for peer instruction (Mazur, 1997; Smith *et al.*, 2009). Materials such as slides, assessment questions, and notes were collected at the meeting, and the group decided to pursue a case study with clicker questions.

After the initial in-person meeting, meetings with faculty across all five institutions were held virtually, recorded for faculty who could not attend, and limited to 1 hour once or twice a semester. A research associate who was a previous faculty partner in the Small World Initiative (www.smallworldinitiative.org) and wanted to engage in DBER projects facilitated the cross-institutional project (author K.N.P.). Her roles were to set the agenda and schedule, distribute materials, solicit feedback, update instructional unit materials based on suggestions, and provide ongoing communication with the faculty by email.

Using the instructional unit in class was not a requirement for participation in the professional development community; of the 16 faculty who helped to develop the instructional unit, eight taught it in class (Supplemental Table 1). The decision by faculty

TABLE 1. Examples of correct, irrelevant/unclear, or incorrect student responses to the AACR questions about how a mutation that results in a premature stop codon affects DNA replication, transcription, and translation

	Correct	Irrelevant/unclear	Incorrect
Replication	"Since during DNA replication only one nucleotide is read at a time, the mutation will have no effect on DNA replication." "Replication will not be altered because it does not deal with the sequence of amino acids." "This will not affect DNA replication."	"Nonsense—incorrectly inserting a stop codon and making the gene stop before the process is complete." "All genes replicated from this gene will end up having the stop codon." "It will also create DNA in the daughter cells with a mutation."	"This nonsense mutation will end replication early." "The DNA replication stops prematurely and creates a nonsense protein that is nonfunctional." "The DNA replication will stop early and will not include the entire DNA sequence. Not all of the DNA will be replicated."
Transcription	"This won't influence transcription because RNA polymerase doesn't read codons." "Transcription won't be affected because stop codons apply to translation." "This alteration does not affect the process of transcription although it affects the transcripted mRNA during translation only."	"Nonsense." "The amino acid that is transcribed would be change to a stop codon." "Mutations such as these can cause genetic disorders. A missense mutation will change the amino acid sequence which can change the function of the protein."	"The transcription of mRNA to amino acids is a process that reads three bases at a time. When the sequence reads over the stop codon, it will prematurely stop the process of transcription." "The RNA sequence will be much shorter." "Since it is now a stop codon it will result in a shorter RNA strand."
Translation	"Translation will be halted prematurely as the ribosome reads the stop codon." "This change would influence translation because the stop codon would end translation early and make the amino acid sequence shorter which will lead to a shorter polypeptide chain." "Translation will end early, resulting in a shorter protein."	"Not enough protein will be produced." "The protein will form incorrectly because it will be missing multiple amino acids from its structure." "Many amino acids will not be made."	"The process of translating DNA to mRNA is one that is performed one base at a time. Although there was an alteration in one base, there will be no harm done to the process of translation, because the stop codon only applies to transcription." "Translation is unaffected by this alteration." "Translation will be affected because the short mRNA strand will make a shorter protein."

to use the instructional unit in their classrooms was driven largely by what courses were being taught in a given semester.

CROSS-INSTITUTIONAL GROUP DYNAMICS

The cross-institutional community provided an opportunity for faculty to discuss a common conceptual difficulty and to develop an instructional unit. During the first virtual meeting, the faculty decided to develop a case study about Duchenne muscular dystrophy, using formative assessment clicker questions and peer instruction (Mazur, 1997; Smith *et al.*, 2009).

In an effort to minimize time investment for the faculty in future meetings, the facilitator sent out instructional materials such as slides ahead of the meeting and solicited suggestions from the faculty. That is, before a virtual meeting, the facilitator received suggested revisions and edited PowerPoint slides. With the intention of protecting faculty time, the facilitator made changes and presented these revised slides at the virtual meeting. She intentionally did not make additional contributions of her own. However, the faculty, who did not see their colleagues' suggestions but only the subsequent changes, assumed that the changes observed were the ideas and opinions of the facilitator. Early reactions from faculty indicated that they felt the facilitator was driving the project. For example, instead of calling it "our project," faculty referred to it as "your project" in reference to the facilitator.

To dispel this perception, all subsequent input sent to the facilitator was shared with the group as anonymous comments to be collectively addressed by the group before changes were made. During the virtual meetings, the faculty discussed the comments and collaborated on how to address the suggestions. The facilitator took notes, clarified recommendations from the faculty, and then made the recommended changes *after* the meeting, providing the notes and actions items to the faculty along with the changed materials. This altered approach seemed to improve the productivity of the discussions about teaching and student learning and reinforce that the instructional unit, including all modifications, resulted from the faculty members' collective work. Notably, removing individual contributions and increasing joint responsibility is an approach that has been shown to contribute to increased cohesion and productivity in group dynamics (Lawler, 2001).

We also had to resolve the issue of faculty skepticism about collaborative lesson design. At the beginning of the project, a subset of faculty participants felt that we were spending too much time on the role of stop codons on DNA replication, transcription, and translation and were hesitant to use class time on a case study approach with several active-learning instructional techniques (i.e., clicker questions, group discussion). We continued to engage these faculty members by sending them data from their own students and aggregate student scores for comparison, asking for their opinions on how to connect the instructional materials to additional concepts in the course and continually inviting them to participate in the virtual meetings and publication of the lesson. All of the faculty remained involved in the group and participated in the publication that describes the final product (Pelletreau *et al.*, 2016).

TABLE 2. Core components of the instructional unit developed and modified by faculty using student data to drive the change[a]

	Version		
	1	2	3
Total slides	20	27	30
Discussion points	2	5	6
Predictions	1	3	3
Animations	0	2	2
Total clicker questions	10	8	9
Clicker questions on:			
Intron and silent mutations affecting phenotype	2	0	0
Missense mutation affecting phenotype	1	1	1
Promoter mutation affecting phenotype	3	2	2
DNA replication	1	1	2
Transcription	1	2	2
Translation	1	1	1
Determining which nucleotide changes are mutations	1	1	1

[a]Discussion points are posed to the class as open-response questions, clicker questions are multiple-choice questions that students discuss with their peers and answer with a clicker, predictions are times when the students are asked to predict outcomes as either clicker or discussion questions, and animations are short animated films developed by the faculty to show the interactions of RNA polymerase and the ribosome with the stop codon. The concepts addressed in the clicker questions are also listed.

THE EFFECTS OF ITERATIVE INSTRUCTIONAL UNIT DEVELOPMENT ON STUDENT LEARNING

In total, three versions of a 50-minute instructional unit (referred to as versions 1, 2, and 3) were developed and assessed by the faculty (Table 2). Three faculty taught version 1 of the instructional unit (Supplemental Table 2), adhering to the same within-semester implementation strategy (Figure 3). Because the faculty expressed concerns about making sure they covered all the content in the central dogma of biology section of their course, they taught DNA replication, transcription, and translation as they normally would, and asked students to answer the AACR stop-codon questions (preassessment). For most faculty, the preassessment was given a few days after the teaching of the concepts of the central dogma of biology, which was 1–2 days before the instructional unit was taught. The preassessment provided faculty the percent of students who had unclear or incorrect understanding of the effects of a stop codon after their instruction. Faculty then taught the instructional unit and asked students to answer the same AACR stop-codon questions 7–10 days after the instructional unit was complete (postassessment). The postassessment provided levels of student understanding of the same concepts after the students participated in the instructional unit.

The faculty group then discussed the student results of the AACR questions generated by version 1 (Figure 4A: aggregate results that include all participating students independent of their institutions; Supplemental Table 3: results by individual instructor). The students showed positive learning gains (Figure 4A), and the majority of students who answered correctly on the preassessment questions continued to answer correctly on the postassessment (Supplemental Table 4). However, because students still scored relatively poorly on the transcription question even after participating in the instructional unit (Figure 4A), the group decided to make revisions (Table 2), including adding more discussion points and new animations. Because the length of the lesson was a limiting factor, the faculty decided to remove some version 1 clicker questions that more than 95% of the students in multiple classrooms answered correctly, such as whether silent mutations are likely causes of Duchenne muscular dystrophy. The faculty also revised existing clicker questions to include more predictions; Supplemental Figure 1 shows an example of how the faculty modified a yes or no question about mRNA length into the prediction of mRNA size on a Northern blot.

Version 2 of the instructional unit was taught in five classes (Supplemental Table 2). Compared with version 1, the version 2 aggregate student learning gains increased (Figure 4B and Supplemental Table 3). The preassessment scores were also higher, likely due to the variation in the courses, instructors, and semester in which the course was taught (Supplemental Table 2). Similar to version 1, the majority of students who answered correctly on the preassessment also answered correctly on the postassessment (Supplemental Table 4).

When the faculty met to talk about the aggregate student data from version 2, they made additional *minor* revisions: they added two overview slides that listed all parts of the central dogma of biology (DNA replication, transcription, and translation) to help orient students, one additional clicker question about DNA replication, and a discussion point about the ribosome recognizing stop codons (Table 2). Version 3 of the instructional unit was subsequently taught in 10 classrooms with the greatest variation in course type (nonmajors biology, majors biology, genetics, molecular and cell biology; Supplemental Table 2). The average preassessment scores were similar to those for version 1, and the learning gains were higher (Figure 4C and Supplemental Table 3). As with versions 1 and 2, the majority of students who answered correctly on the preassessment also answered correctly on the postassessment (Supplemental Table 4).

To address the concern that the gains in student performance were due to repeated exposure to the question, also known as the practice effect (Wing, 1980; O'Neill *et al.*, 2015), an instructor (Riley) who was coteaching with one of the faculty participants asked the AACR stop-codon questions before and after her central dogma of biology unit but did *not* teach the instructional unit (Supplemental Table 2). In her

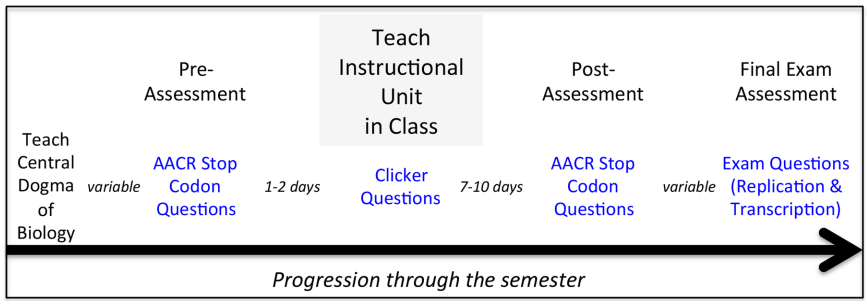

FIGURE 3. Implementation of the instructional unit and data collected (noted in blue font) at each time point. This timeline was used for all versions of the instructional unit.

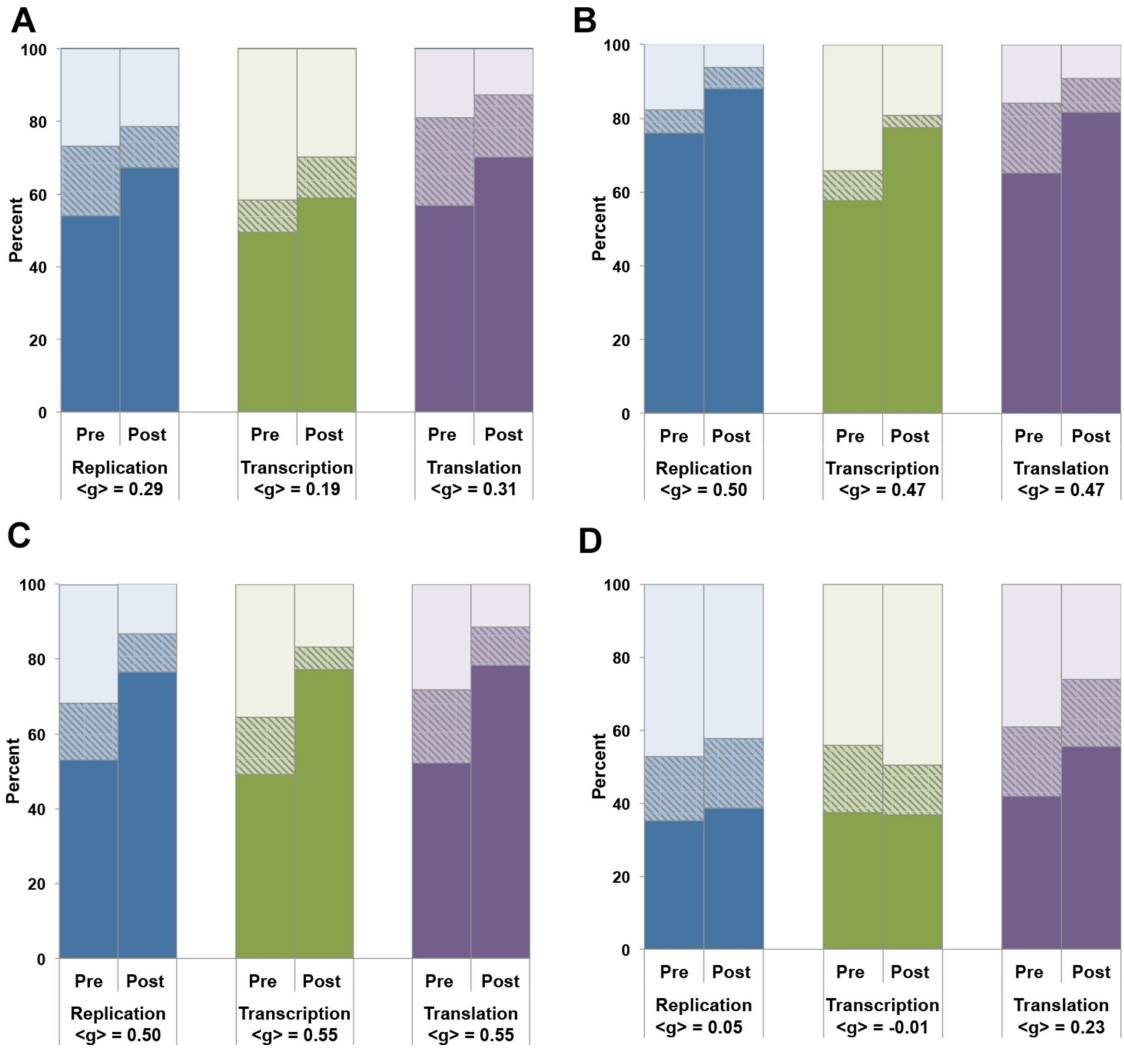

FIGURE 4. Aggregate student performance on the AACR stop-codon questions before (preassessment) and after (postassessment) for (A) version 1, (B) version 2, and (C) version 3 of the instructional unit and (D) a control group in which students answered the questions twice (1 week apart) without participating in the instructional unit. Student responses are binned as correct (dark color), irrelevant/unclear (hashed color), or incorrect (light color) for questions on replication (blue), transcription (green), and translation (purple). Normalized learning gains <g> (Hake, 1998) for students who answered the question correctly are presented below each category [<g> = (% of students who scored correct on the posttest) – (% of students who scored correct on the pretest)/(100 – % of students who scored correct on the pretest)]. Different instructors and courses from five institutions taught each version; course details are described in Supplemental Table 2. Supplemental Table 3 shows percent correct and learning gains for each instructor.

classroom, there were minimal to no positive learning gains for the DNA replication question and the transcription question, and a modest positive learning gain for the translation question (Figure 4D). These results suggest that students benefit from engaging in an activity that explores whether stop codons have a role in DNA replication and transcription, and the learning gains are not due to repeated exposure to a question.

The majority of the faculty also measured student performance on one or two final exam questions that were administered at multiple institutions and saw higher performance from students who participated in versions 2 and 3 of the instructional unit, compared with students who participated in version 1 (Figure 5, exam questions shown in Supplemental Figure 2). Performance on the DNA replication final exam question was similar for versions 2 and 3 (Figure 5A). For the transcription question, student performance was also similar for versions 2 and 3, with the exception of one low-performing class for version 3 (Figure 5B). Taken together, these results show a link between the use of data to iteratively revise an instructional unit and subsequent improvements in student learning. The faculty collectively coauthored a paper for the journal *CourseSource* that provides the instructional materials, lesson timeline, and implementation guidelines for version 3 (Pelletreau *et al.*, 2016).

THE EFFECTS OF FACULTY PROFESSIONAL DEVELOPMENT ON INSTRUCTIONAL PRACTICES

To measure how faculty participation impacted teaching practices, faculty members were observed using the Classroom

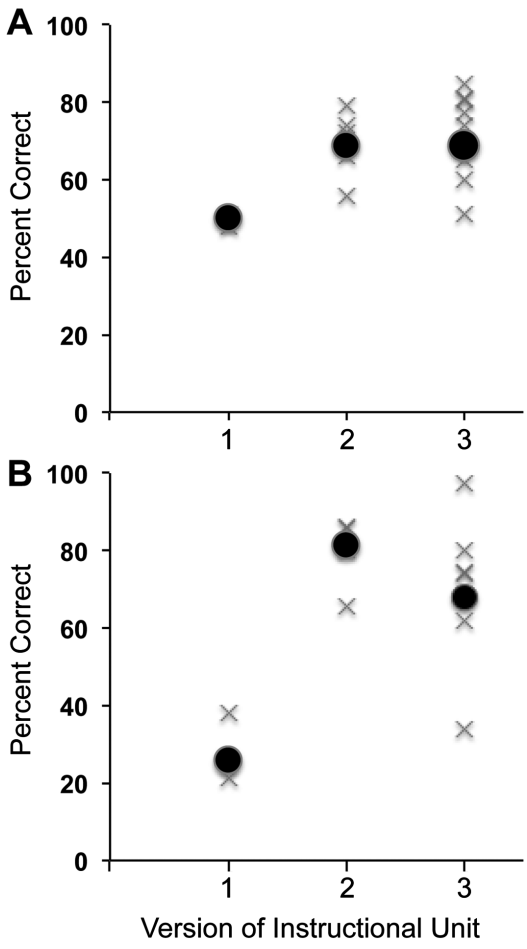

FIGURE 5. Student performance on two shared final exam questions given after each version of the instructional unit, one on replication (A) and the other on transcription (B). Average student scores from each class that used the exam questions are denoted by gray X's; filled circles represent the weighted average for all classes to account for different class sizes. Total number of students answering each question is as follows: DNA replication version 1 (n = 948), version 2 (n = 777), and version 3 (n = 2560); transcription version 1 (n = 948), version 2 (n = 629), and version 3 (n = 1689). The student numbers vary because some instructors chose to only ask one of the two questions on their exams.

Observation Protocol for Undergraduate STEM (COPUS; Smith *et al.*, 2013), adapted from the Teaching Dimensions Observation Protocol (Hora *et al.*, 2013; Hora and Ferrare, 2014). This observation protocol uses a series of codes to characterize instructor and student practices in the classroom in each 2-minute interval throughout the duration of a class period (Smith *et al.*, 2013, 2014). Because faculty members were teaching at different institutions over several semesters, 29 observers were trained in an online training session and then separately coded a set of 20-minute training videos. We then compared observers' pairwise scores and, once they achieved a Cohen's kappa score ≥0.80 (Landis and Koch, 1977), the observers were able to independently collect data in the classrooms.

Faculty were observed on multiple occasions each semester: on days when they taught the instructional unit and on two or more other days of instruction. We used four collapsed instructor COPUS code categories that broadly reflect different activities in the classroom: presenting (e.g., lecturing, real-time writing), guiding (e.g., posing clicker questions, moving and guiding throughout the class), administration, and other (Smith *et al.*, 2014). The codes that make up the guiding collapsed code are more common in active-learning classrooms (Smith *et al.*, 2014). The percent code was calculated by adding the total number of times a code of interest was selected and dividing by the total number of codes observed.

We observed a shift in teaching practices on the day the faculty taught the instructional unit. On days when the faculty were not teaching the instructional unit, they collectively employed a range of instructional practices. For example, the presenting collapsed code comprised <5% to >90% of the codes (Figure 6A). However, on the day these same instructors used any version of the instructional unit, there was an overall shift in instructional practices; the maximum percent of the presenting collapsed code dropped from >90% to <67% (Figure 6B). During the teaching of the instructional unit, this decrease in the presenting collapsed code was mirrored by an increase in instructor guiding collapsed code.

For the majority of faculty, the instructional unit supported this shift in instructional practices toward more active-learning instructional techniques (Supplemental Figure 3 shows the observation data reorganized by instructor). However, two exceptions occurred. The first was when faculty members were already employing predominantly active-learning strategies (Jackson version 1, Stella versions 2 and 3, and Alex version 3). The second occurred for Lily version 2, who reported feeling rushed that day and implementing the instructional unit in a shortened time at the end of the class period. For the remaining instructors, teaching the instructional unit provided them with an opportunity to try new teaching practices using an instructional unit they helped to create.

FACULTY EXPERIENCES

While version 2 of the instructional unit was being implemented, author J.S.M., a postdoctoral researcher who did not participate in the instructional unit development process, used semistructured interviews to elicit feedback from the faculty members (Patton, 2014). The interview questions about the instructional unit were part of a broader interview protocol that examined multiple components of the larger AACR project (McCourt *et al.*, 2017). The interviewer used a list of predetermined questions in addition to follow-up questions to learn more about each faculty member's ideas regarding his or her involvement in the AACR research project and this instructional unit development group. The interviews were transcribed and reviewed to gain insight into faculty perceptions about what types of support are needed to promote instructional change.

The interviews revealed that faculty found participation in this cross-institutional faculty group rewarding, with all stating they would be interested in continuing to work on curriculum development projects together. Multiple faculty mentioned the opportunity to be part of a diverse collaborative group focused on teaching, one in which contributions were valued. The faculty also appreciated the efficiency of the process, made possible by the facilitator who coordinated the implementation process (e.g., sent reminders about giving the pre- and

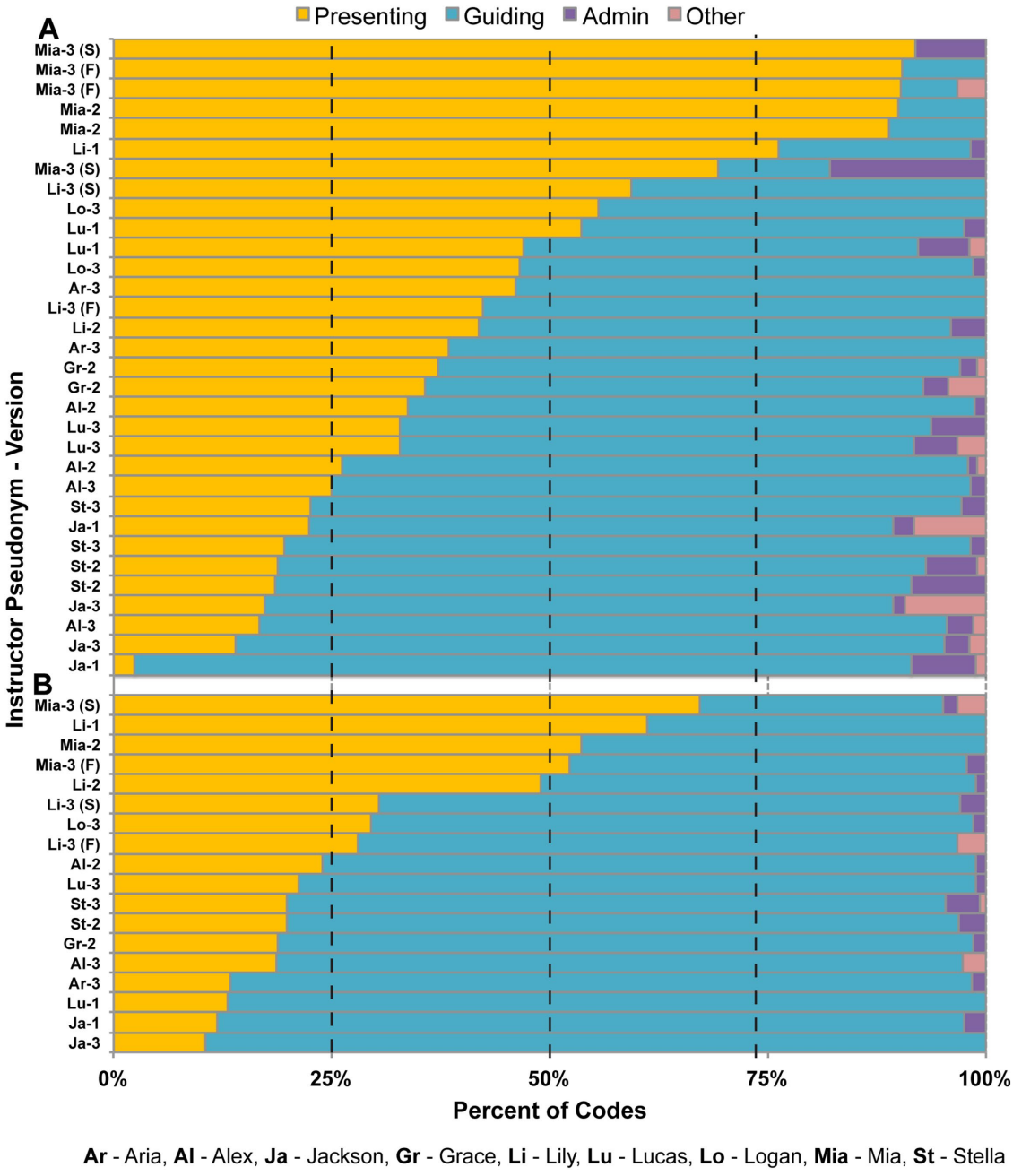

FIGURE 6. Collapsed code COPUS data for the faculty on (A) days when they did *not* teach the instructional unit and (B) days when the instructional unit was taught. Initials correspond to the pseudonym for each faculty member (see figure key), and the number indicates which version of the instructional unit was being used that semester (Supplemental Table 2). The semester is noted after the instructor's name if the same instructor taught version 3 more than once (F, Fall; S, Spring). Because of scheduling conflicts, one faculty member (Lily) had only one observation per semester for a day when she did not teach the instructional unit.

postassessment and exam questions, collected and aggregated data from multiple institutions, organized the virtual meetings, updated the instructional unit, and sent revised versions of the instructional unit to the group for further input). Finally, multiple faculty positively discussed the efficacy of the instructional unit itself, noting that students benefited from synthesizing the information. Illustrative quotes aligned with each of these points are included in Table 3.

TAKEAWAY LESSONS

On the basis of our experience with this faculty professional development model, we offer the following guidelines for departmental chairs, education researchers, professional/educational developers, and others who are working with faculty and are interested in improving student learning and exploring active-learning instructional practices. These faculty groups could be brought together for a variety of reasons,

TABLE 3. Information about the instructional unit development process taken from the faculty interviews

Theme from faculty interview	Illustrative quote
There are benefits to being part of a faculty group with different perspectives	"It's always interesting to see other people's perspectives. You know, sometimes you think you know something and you realize well maybe not or maybe there's a different way to think about it."
	"In the development of that project, there were a whole lot of people contributing with some really good comments. I don't think any [faculty member] stood out any more than anyone else in that group, but that's what's so nice. In the end, it doesn't really matter. It was the fact that all of us, in fact, were contributing that made it such a nice activity."
Efficiency was an important part of the process	"Sometimes you have too many people and it just gets too many opinions, but in this case, you had one person doing stuff [the facilitator] and then they were getting feedback from other people, and that seemed to work pretty well."
	"Cost effective is how I look at it. My participation saves me a huge amount of time [rather] than trying to develop that on my own, so it's a very efficient way to get stuff done."
Faculty were proud of the instructional unit	"Well, I thought the activity we were working toward was going to be really good and it ended up being really good. It was amazingly comprehensive and hit all the points that we would want to hit in a lecture on that topic and, making the students do it all, that was really great."

including common student conceptual difficulties, courses within a department that have shared learning goals, and/or courses with a high failure rate.

1. Take a data-driven iterative approach. Despite calls to use a "culture of evidence" when teaching at the undergraduate level, recent work shows that STEM faculty often rely on their intuition, personal experience, and recommendations from colleagues rather than student learning data when making decisions (Andrews and Lemons, 2015; Marbach-Ad and Hunt Rietschel, 2016; Hora *et al.*, 2017). In the project described here, faculty benefited from a community in which they had opportunities to look at their own students' data, discuss aggregate student data from multiple classrooms, and respond to one another's suggestions. Although previous work has shown that an FLC focused on active learning can lead to measured improvements in student performance in biology courses (Elliott *et al.*, 2016), the work of this faculty group provides evidence that having faculty work together can lead to modifications of instructional materials that improve student learning gains (Figures 4 and 5). Notably, the community that developed around this project was distinct from the faculty members' own institutional cultures and may have resulted in unique patterns of collaboration. We are currently analyzing the audio recordings from the meetings to learn more about how the faculty worked together.
2. Use a designated, knowledgeable facilitator. A facilitator who can minimize barriers (Silverthorn *et al.*, 2006; Henderson and Dancy, 2007; Ortquist-Ahrens and Torosyan, 2009; Brownell and Tanner, 2012; Corbo *et al.*, 2016) and coach the group to continue working together proved vital for this group of faculty. In our project, the facilitator, who was knowledgeable about biology education research issues, reduced faculty time commitment by organizing meetings; coordinating the implementation of the instructional unit; gathering, analyzing, and presenting aggregate data; and making updates to the shared work (e.g., class slides, assessment questions, animations). In addition, the facilitator reinforced faculty ownership by requesting specific feedback, assembling ideas, and stimulating group discussion about these ideas before making any changes to the instructional unit. Importantly, a facilitator can also organize the collaborative publication of the materials when faculty participants may lack the time and/or experience to do so (e.g., Pelletreau *et al.*, 2016).

   For projects that do not have resources to employ a facilitator, faculty could rotate leadership roles among members of the group. Because a facilitator does not necessarily need to be an advanced education researcher, graduate students and postdocs could be involved in making changes to instructional materials, organizing student learning data, and drafting manuscripts as part of their pedagogical training. It is also worth investigating the resources at centers for teaching and learning, where people with an instructional development background are often available for consultations or facilitation.
3. Minimize risk and maximize reward for the faculty. It is important to implement an instructional unit that is likely to be successful. In our program, faculty were asked to teach a active-learning instructional unit that had been developed, organized, discussed, and evaluated by their peers. Although this type of instructional practice was new to some faculty (Figure 6 and Supplemental Figure 3), the possibility of failure was minimized, because the iterative development process resulted in well-vetted materials and the faculty felt comfortable trying a new active-learning approach in their classrooms. For a few faculty, implementing such a unit did not dramatically change their classroom environments, but for others, implementation shifted their classroom environments toward more active-learning practices (Figure 6 and Supplemental Figure 3). Following the holistic changes in instructional practices of these faculty over time and on different lessons will be important to determine whether short-term instructional practice changes persist into long-term changes.

   Involvement in our professional development group also provided faculty coauthorship on a manuscript for their curricula vitae. From the beginning of the collaboration, publication of the lesson was presented to the faculty as an opportunity. Reminding faculty of the goal of collaboratively publishing the instructional unit helped encourage their participation in editing the instructional unit, attending meetings, and collecting and sharing student performance data.

4. A little variation in teaching practices is okay. Although a single set of slides and clicker questions was developed by and disseminated to the faculty, there was still variation in implementation (Figure 6 and Supplemental Figure 3). This outcome is consistent with prior work showing that few faculty fully adopt curricular materials and instead often make changes (Henderson and Dancy, 2008). Fidelity of implementation is known to potentially affect outcomes (Turpen and Finkelstein, 2009; Daubenmire *et al.*, 2015; Dancy *et al.*, 2016; Stains and Vickrey, 2017). However, regardless of a faculty member's unique instructional fingerprint, student learning improved from pre- to postassessment in all classes (Supplemental Table 3). Thus, future studies will further explore links between individual instructional practices and student learning outcomes.

CONCLUSIONS

This professional development opportunity brought together faculty from multiple institutions with a shared interest in improving student learning and instruction of complex concepts critical to their courses. We found that the faculty used student learning data to motivate the development of the instructional unit and iteratively improve their teaching, a facilitator minimized faculty time investment while maintaining a sense of ownership, and a collaborative publication was strong incentive for faculty. The instructional unit not only improved student learning of DNA replication, transcription, and translation, but also gave faculty an opportunity to try new active-learning instructional strategies in their classrooms, potentially leading to increased adoption of such practices. This approach to professional development had positive outcomes for students and faculty alike and can be applied to settings where faculty members are geographically remote but share a common pedagogical interest. Furthermore, the published products of these groups (e.g., Pelletreau *et al.*, 2016) benefit the faculty participants and can be shared broadly, amplifying the effect this professional development program can have on transforming STEM education.

ACKNOWLEDGMENTS

This material is based on work supported by the NSF under grant numbers DUE 1347578 and 1322851. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. We thank the instructors who participated in this study and the following individuals for providing feedback on the essay: Carolyn Aslan, Natasha Holmes, Peter LePage, Kelly Nielson, Melissa Pirkey, and Lisa Sanfilippo.

REFERENCES

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher*, *41*(1), 16–25. doi: 10.3102/0013189X11428813

Anderson, W. A., Banerjee, U., Drennan, C. L., Elgin, S. C. R., Epstein, I. R., Handelsman, J., ... Warner, I. M. (2011). Changing the culture of science education at research universities. *Science*, *331*(6014), 152–153. doi: org/10.1126/science.1198280

Andrews, T. C., & Lemons, P. P. (2015). It's personal: Biology instructors prioritize personal evidence over empirical evidence in teaching decisions. *CBE—Life Science Education*, *14*(1), ar7. doi: org/10.1187/cbe.14-05-0084

Beach, A. L., & Cox, M. D. (2009). The impact of faculty learning communities on teaching and learning. *Learning Communities Journal*, *1*(1), 7–27.

Brownell, S. E., & Tanner, K. D. (2012). Barriers to faculty pedagogical change: Lack of training, time, incentives, and ... tensions with professional identity? *CBE—Life Science Education*, *11*(4), 339–346. doi: org/10.1187/cbe.12-09-0163

Corbo, J. C., Reinholz, D. L., Dancy, M. H., & Deetz, S. (2016). Framework for transforming departmental culture to support educational innovation. *Physical Review Physics Education Research*, *12*. doi: org/10/1103/PhysRevPhysEducRes.12.010113

Cox, M. D. (2004). Introduction to faculty learning communities. *New Directions for Teaching and Learning*, *97,* 5–23.

Cox, M. D. (2016). Four positions of leadership in planning, implementing, and sustaining faculty learning community programs. *New Directions for Teaching and Learning*, *148,* 85–96. doi: org/10.1002/tl.20212

Dancy, M., Henderson, C., & Turpen, C. (2016). How faculty learn about and implement research-based instructional strategies: The case of peer instruction. *Physical Review Physics Education Research*, *12*(1). doi: org/10.1103/PhysRevPhysEducRes.12.010110

Daubenmire, P. L., Bunce, D. M., Draus, C., & Frazier, M. (2015). During POGIL implementation the professor still makes a difference. *Journal of College Science Teaching*, *44*(5), 72–81.

Derting, T. L., Ebert-May, D., Henkel, T. P., Maher, J. M., Arnold, B., & Passmore, H. A. (2016). Assessing faculty professional development in STEM higher education: Sustainability of outcomes. *Science Advances*, *2*(3). doi: 10.1126/sciadv.1501422

Ebert-May, D., Derting, T. L., Henkel, T. P., Maher, J. M., Momsen, J. L., Arnold, B., & Passmore, H. A. (2015). Breaking the cycle: Future faculty begin teaching with learner-centered strategies after professional development. *CBE—Life Sciences Education*, *14*(2), ar22. doi: 10.1187/cbe.14-12-0222

Ebert-May, D., Derting, T. L., Hodder, J., Momsen, J. L., Long, T. M., & Jardeleza, S. E. (2011). What we say is not what we do: Effective evaluation of faculty professional development programs. *BioScience*, *61*(7), 550–558. doi: 10.1525/bio.2011.61.7.9

Elliott, E. R., Reason, R. D., Coffman, C. R., Gangloff, E. J., Raker, J. R., Powell-Coffman, J. A., & Ogilvie, C. A. (2016). Improved student learning through a faculty learning community: How faculty collaboration transformed a large-enrollment course from lecture to student centered. *CBE—Life Science Education*, *15*(2), ar22. doi: 10.1187/cbe.14-07-0112

Emtage, L., Bradbury, L., Coleman, N., Devenport, D., Nietzel, A., & Grew, J. (2016). Cell signaling pathways: A case study approach. *CourseSource*, *3,* 1–7. doi: 10.24918/cs.2016.9

Freeman, P. L., Maki, J. A., Thoemke, K. R., Lamm, M. H., & Coffman, C. R. (2017). Evaluating the quick fix: Weight loss drugs and cellular respiration. *CourseSource*. doi: 10.24918/cs.2017.17

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences USA*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, *66*(1), 64. doi: 10.1119/1.18809

Haudek, K. C., Kaplan, J. J., Knight, J., Long, T., Merrill, J., Munn, A., ... Urban-Lurain, M. (2011). Harnessing technology to improve formative assessment of student conceptions in STEM: Forging a national network. *CBE—Life Science Education*, *10*(2), 149–155. doi: 10.1187/cbe.11-03-0019

Haudek, K. C., Moscarella, R. A., Weston, M., Merrill, J., & Urban-Lurain, M. (2015). Construction of rubrics to evaluate content in students' scientific explanation using computerized text analysis. *National Association for Research in Science Teaching conference* (pp. 1–30).

Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *Journal of Research in Science Teaching*, *48*(8), 952–984. doi: 10.1002/tea.20439

Henderson, C., & Dancy, M. H. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical Review Special Topics—Physics Education Research*, *3*(2), 010103. doi: 10.1103/PhysRevSTPER.3.020102.

Henderson, C., & Dancy, M. H. (2008). Physics faculty and educational researchers: Divergent expectations as barriers to the diffusion of innovations. *American Journal of Physics*, *76*(1), 79–91. doi: 10.1119/1.2800352

Hora, M. T., Bouwma-Gearhart, J., & Park, H. J. (2017). Data driven decision-making in the era of accountability: Fostering faculty data cultures for learning. *Review of Higher Education*, *40*(3), 391–426.

Hora, M. T., & Ferrare, J. J. (2014). Remeasuring postsecondary teaching: How singular categories of instruction obscure the multiple dimensions of classroom practice. *Journal of College Science Teaching*, *43*(3), 36–41.

Hora, M. T., Oleson, A., & Ferrare, J. J. (2013). *Teaching dimensions observation protocol (TDOP) user's manual*. Madison: Wisconsin Center for Education Research, University of Wisconsin–Madison. Retrieved August 7, 2017, from https://tdop.wceruw.org/Document/TDOP-2.1-Users-Guide.pdf

Hoskinson, A. M., Conner, L., Leigh, M. B., Martin, A. P., & Powers, T. (2014). Coevolution or not? Crossbills, squirrels and pinecones. *CourseSource*, *1,* 1–8. doi: 10.24918/cs.2014.4.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lawler, E. J. (2001). An affect theory of social exchange. *American Journal of Sociology*, *107*(2), 321–352.

Marbach-Ad, G., & Hunt Rietschel, C. (2016). A case study documenting the process by which biology instructors transition from teacher-centered to learner-centered teaching. *CBE—Life Science Education*, *15*(4), ar62. doi: 10.1187/cbe.16-06-0196

Mazur, E. (1997). *Peer instruction: A user's manual*. Upper Saddle River, NJ: Pearson Higher Education.

McCourt, J. S., Andrews, T. C., Knight, J. K., Merrill, J. E., Nehm, R. H., Pelletreau, K. N., … Lemons, P. P. (2017). What motivates biology instructors to engage and persist in teaching professional development. *CBE—Life Science Education*, *16,* ar54. doi: 10.1187/cbe.16-08-0241

Moharreri, K., Ha, M., & Nehm, R. H. (2014). EvoGrader: An online formative assessment tool for automatically evaluating written evolutionary explanations. *Evolution: Education and Outreach*, *7*(1), 15.

O'Neill, T. R., Sun, S., Peabody, M. R., & Royal, K. D. (2015). The impact of repeated exposure to items. *Teaching and Learning in Medicine*, *27*(4), 404–409. doi: 10.1080/10401334.2015.1077131

Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations for science assessment design: Knowing what students know about evolution. *Journal of Research in Science Teaching*, *49*(6), 744–777.

Ortquist-Ahrens, L., & Torosyan, R. (2009). The role of the facilitator in faculty learning communities: Paving the way for growth, productivity, and collegiality. *Learning Communities Journal*, *1*(1), 29–62.

Patton, M. Q. (2014). *Qualitative evaluation and research methods* (2nd ed.). Newbury Park, CA: Sage.

Pelletreau, K. N., Andrews, T. C., Armstrong, N., Bedell, M. A., Dastoor, F., Dean, N., … Smith, M. K. (2016). A clicker-based case study that untangles student thinking about the processes in the central dogma. *CourseSource*, *3,* 1–10. doi: 10.24918/cs.2016.15

Penberthy, D. L., & Millar, S. B. (2002). The "hand-off" as a flawed approach to disseminating innovation: Lessons from chemistry. *Innovative Higher Education*, *26*(4), 251–270. doi: 10.1023/A:1015828913383

Pfund, C., Miller, S., Brenner, K., Bruns, P., Chang, A., Ebert-May, D., … Handelsman, J. (2009). Summer institute to improve university science teaching. *Science*, *324*(5926), 470–471. doi: 10.1126/science.1170015

Prevost, L. B., Smith, M. K., & Knight, J. K. (2016). Using student writing and lexical analysis to reveal student thinking about the role of stop codons in the central dogma. *CBE—Life Science Education*, *15*(4), ar65. doi: 10.1187/cbe.15-12-0267

Richlin, L., & Cox, M. D. (2004). Developing scholarly teaching and the scholarship of teaching and learning through faculty learning communities. *New Directions in Teaching and Learning*, *97,* 127–135. doi: 10.1002/tl.139.

Sestero, C., Tinsley, H., Ye, Z.-H., Zhang, X., Graze, R., & Kearley, M. (2014). Using the cell engineer/detective approach to explore cell structure and function. *CourseSource*, *1,* 1–5. doi: 10.24918/cs.2014.7

Sharp, S., & McLaughlin, P. (1997). Disseminating development initiatives in British higher education: A case study. *Higher Education*, *38*(3), 309–329. doi: 10.1023/A:1002959730812

Silverthorn, D. U., Thorn, P. M., & Svinicki, M. D. (2006). It's difficult to change the way we teach: Lessons from the Integrative Themes in Physiology curriculum module project. *Advances in Physiology Education*, *30*(4), 204–214. doi: 10.1152/advan.00064.2006

Smith, M. K., Jones, F. H. M., Gilbert, S. L., & Wieman, C. E. (2013). The Classroom Observation Protocol for Undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices. *CBE—Life Science Education*, *12*(4), 618–627. doi: 10.1187/cbe.13-08-0154

Smith, M. K., & Knight, J. K. (2012). Using the Genetics Concept Assessment to document persistent conceptual difficulties in undergraduate genetics courses. *Genetics*, *191,* 21–32.

Smith, M. K., Vinson, E. L., Smith, J. A., Lewin, J. D., & Stetzer, M. R. (2014). A campus-wide study of STEM courses: New perspectives on teaching practices and perceptions. *CBE—Life Sciences Education*, *13*(4), 624–635. doi: 10.1187/cbe.14-06-0108

Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Su, T. T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science*, *323*(5910), 122–124. doi: 10.1126/science.1165919

Smith, M. K., Wood, W. B., & Knight, J. K. (2008). The Genetics Concept Assessment: A new concept inventory for gauging student understanding of genetics. *CBE—Life Science Education*, *7*(4), 422–430.

Stains, M., & Vickrey, T. (2017). Fidelity of implementation: An overlooked yet critical construct to establish effectiveness of evidence-based instructional practices. *CBE—Life Science Education*, *16*(1), rm1. doi: 10.1187/cbe.16-03-0113.

Thompson, K. V., Marbach-Ad, G., Egan, L., & Smith, A. C. (2015). Faculty learning communities: A professional development model that fosters individual, departmental and institutional impact. In Weaver, G. C., Burgess, W. D., Childress, A. L., & Slakey, L. (Eds.), *Transforming institutions: Undergraduate STEM education for the 21st century* (pp. 312–324). West Lafayette, IN: Purdue University Press.

Turpen, C., & Finkelstein, N. D. (2009). The construction of different classroom norms during Peer Instruction: Students perceive differences. *Physical Review Physics Education Research*, *6*(2), 1–22. doi: 10.1103/PhysRevSTPER.6.020123

Wieman, C., Perkins, K., & Gilbert, S. (2010). Transforming science education at large research universities: A case study in progress. *Change: The Magazine of Higher Learning*, *42*(2), 7–14.

Wing, H. (1980). Practice effects with traditional mental test items. *Applied Psychological Measurement*, *4*(2), 141–155.

Wood, W. B., & Handelsman, J. (2004). Meeting report: The 2004 National Academies Summer Institute on Undergraduate Education in Biology. *Cell Biology Education*, *3*(4), 215–217. doi: 10.1187/cbe.04-07-0057

Wright, L. K., Fisk, J. N., & Newman, D. L. (2014). DNA→RNA: What do students think the arrow means? *CBE—Life Science Education*, *13*(2), 338–348.