

1-1-2016

# Stirring the Pot: Supporting and Challenging General Education Science, Technology, Engineering, and Mathematics Faculty to Change Teaching and Assessment Practice

Vicki Stieha

*Boise State University*

Susan E. Shadle

*Boise State University*

Sharon Paterson

*Boise State University*

# **Stirring the Pot: Supporting and Challenging General Education Science, Technology, Engineering, and Mathematics Faculty to Change Teaching and Assessment Practice**

**Vicki Stieha**

Boise State University

**Susan E. Shadle**

Boise State University

and

**Sharon Paterson**

Boise State University

## **Abstract**

Evidence-based instructional practices (EBIPs) have been associated with positive student outcomes; however, institutions struggle to catalyze widespread adoption of these practices in general education science, technology, engineering, and mathematics (STEM) courses. Further, linking EBIPs with integrated learning assessment is rarely discussed in the literature, even though principles of continuous course design for quality higher education assume the connection of learning outcomes, teaching practices, and assessment. This qualitative action research study documents faculty attempting to utilize EBIPs and authentic assessment when they were provided support and accountability over multiple semesters. We document participants' current practices and encourage shifts in both teaching and assessment practices targeting greater student success in STEM general education courses. Narrative data drawn from interviews and written reflection describe the impact of structured faculty observation and cross-disciplinary conversation on participants' pedagogical and assessment choices. The faculty participants' voices vividly illustrate our findings: incentives, reflection, mentorship, and collaboration over multiple semesters support changing pedagogical practices and integrating outcomes assessment. While STEM-based, the study findings are applicable across the general education curriculum.

**Keywords:** evidence-based instructional practices, faculty development, learning outcomes assessment, STEM education, student success, teaching observation

In the United States, the number of students who start and do not complete a science, technology, engineering, and mathematics (STEM) undergraduate degree pathway ranges from 40 to 60 percent (Chen & Soldner, 2013; Seymour & Hewitt, 1997). National calls to increase the number of STEM graduates such as "Engage to Excel" (Olson & Riordan, 2012) assert the need for, among other changes, transforming the way courses are taught in order to improve these outcomes. What we rarely discuss is that the pathway to STEM degrees often begins in general education science and mathematics courses. Effective teaching and assessment of the skills and knowledge gained in general education STEM courses is important whether or not students remain in a STEM major, as a "scientifically literate citizenry" is an essential component of a liberal education (Association of American Colleges and Universities [AAC&U], 2014). Furthermore, it is a challenge to have general education STEM courses satisfy content requirements and meet general education outcomes. Therefore, enhancing the effectiveness of our teaching methods and improving the learning outcomes in general education science and mathematics classes benefit all undergraduates, whether or not they ultimately pursue STEM majors.

With that background in mind, this study is the result of one institution's efforts to help faculty understand and practice the connections between essential elements of teaching, learning, and assessment. A midsize doctoral research institution with a metropolitan mission, the university works to intentionally attend to pedagogy and student learning

through its strong scholarship of teaching and learning community. In this article we document the change of practice that happens when there is support and accountability for faculty to adopt connected active learning and assessment practices over multiple semesters. Our findings point to the importance of incentives for faculty to engage in change efforts, for them to have opportunities to engage and reflect with other faculty who are also working to revise their teaching and assessment practices, and for them to have ongoing contact with knowledgeable mentors for active learning and assessment practices. The result of this study surfaces tensions in reforming teaching and assessment practices and highlights the networks that may be formed during teaching reform processes, raising prospects for continuing educational change via such networks.

While the study did not begin with the intention to “change” faculty teaching practice, ultimately that was the direction it took. Our initial intention, as discussed in the methods section, was to document existing teaching and assessment practices in general education STEM courses on our campus. These science and mathematics courses, which may be taken to fulfill general education requirements by any student for any major, are specified as required courses for STEM degrees. It was the participants’ insights that shifted our efforts from mere observing to catalyzing change, or “stirring the pot,” by adding faculty development activities, project mentors, and collaborative discussions following the study activities each year. Thus, the act of participating in the project fostered changes in teaching and assessment. Prior to detailing the phases of the action research, we situate our findings in the literature of barriers and drivers to the adoption of evidence-based teaching in STEM disciplines. In addition, we discuss pertinent scholarship about learning outcomes assessment in this segment of the college curriculum.

### Review of Literature

Because the literature that directly connects evidence of learning with evidence-based instructional practices (EBIPs) is scant, many studies point to proxies for student learning and discuss student success in general, such as students’ academic progress (e.g., credits earned, grades) and engagement (Kuh, Kinzie, Buckley, Bridges, & Hayek, 2006). Disciplinary-based studies in physics (Hake, 1998; Watkins & Mazur, 2013), biological science (Eastwood, Sadler, Sherwood, & Schlegel, 2013; Kazempour, Amirshokoohi, & Harwood, 2012), chemistry (Carmel, Jessa, & Yeziarski, 2015; Murphy, Picione, & Holme, 2010), mathematics (Kogan & Laursen, 2014; Love, Hodge, Grandgenett, & Swift, 2014), and engineering (Olakanmi & Doyoyo, 2014; Pakala & Bose, 2015) provide compelling evidence that active learning pedagogies have a positive impact on learners’ conceptual understandings. An often cited meta-analysis concludes that traditional lecturing in STEM courses yields a 1.5 times higher course failure rate than active learning methods (Freeman et al., 2014).

Regardless of this body of evidence, most STEM faculty have not shifted to embrace EBIPs or active learning (Berrett, 2012; Brownell & Tanner, 2015), and numerous studies document barriers for faculty to significantly revise teaching practices. The most common barriers cited are lack of time, training, and incentives (Andrews & Lemons, 2015; Brownell & Tanner, 2015; Henderson, Beach, & Finkelstein, 2011); felt tension in faculty identity as teacher versus researcher (Brownell & Tanner, 2015); and a lower sense of self-efficacy when attempting unfamiliar instructional strategies (Chang, Lin, & Song, 2011). The most common methods used by faculty developers and others to disseminate EBIPs to faculty lack “coordinated and focused efforts lasting over a period of time,” “the use of performance evaluation and feedback,” and a “deliberate focus on changing faculty conceptions”—all essential elements to reinforce adoption (Henderson et al., 2011, p. 972).

Despite the variables that tend to detract from change efforts, there are strategies and conditions that bolster pedagogical change. For example, department cultures that are learning-centered (Dubrow, 2004; Gess-Newsome, Southerland, Johnston, & Woodbury, 2003; Rutz, Condon, Iverson, Manduca, & Willett, 2012) and those encouraging innovation and risk are associated with student learning gains (Rutz et al., 2012). Systematic faculty observation can also support a culture of active learning on a campus (Fox & Hackerman, 2003; Galbraith & Merrill, 2012; Gasiewski, Eagan, Garcia, Hurtado, & Chang, 2012).

A recent proliferation of teaching observation instruments provides for constructive feedback on practice for college-level faculty. These instruments focus attention on observable behaviors and standardize feedback, contrasting with systems that may otherwise be subjective and inconsistent (Centra, 1994; Martinez, Borko, Stecher, Luskin, & Klower, 2012). The Reformed Teaching Observation Protocol (Piburn et al., 2000; Sawada et al., 2002), PORTAAL (Eddy, Coverse, & Wenderoth, 2015), EQUIP (Marshall, Horton, & White, 2009), and Classroom Observation Protocol for

Undergraduate STEM (Smith, Jones, Gilbert, & Wieman, 2013) instruments are just a few that support systematically gathering evidence of active learning.

Integrating EBIPs into teaching practice is facilitated through action research and faculty reflection (Henderson et al., 2011), as well as combining faculty development and observation with institutional and departmental supports (Rutz et al., 2012; Sellheim & Weddle, 2015). While not STEM-focused, Rutz et al. (2012) present some evidence linking faculty development focused on EBIPs and student gains in critical-thinking learning outcomes, a connection that has been elusive. This literature suggests that connecting teaching to course-based authentic learning assessment provides the best opportunity to build a cycle of continuous course design improvement, enhancing student learning outcomes (Banta & Blaich, 2011; Biggs & Teng, 2011; Fink, 2003; Rutz et al., 2012).

Authentic assessment contributes to continuous course improvement, considers student performance relative to target learning outcomes established for a course, and is conducted in a manner consistent with the realistic application of knowledge. Wiggins (1993) is credited with establishing this notion of assessment in contrast to norm- and criterion-referenced surveys and exams often removed from the performance or application of knowledge. Challenges to authentic, or learning-centered, assessment include many of the same challenges encountered with learner-centered pedagogies or EBIPs. Just a few of these challenges include heavy reliance on grades as a proxy for learning (Wiggins, 1993), reliance upon student evaluations for teacher effectiveness (Fox & Hackerman, 2003), low self-efficacy in assessment design skills (Chang et al., 2011), and concerns about time to design and score (Webber, 2012). Yet it is widely agreed that authentic assessment is likely to support a primary goal of assessment—enhancing student learning by informing teaching practice and clarifying learning achievement for the student (Heywood, 2000; Webber, 2012; Wiggins, 1993).

All too often assessment efforts fall short of actually generating evidence that is employed to bring about educational change (Graff, 2010; Grassian, 2013). For example, only 6 percent of the profiles of good assessment practices reviewed by Banta, Jones, and Black in *Designing Effective Assessment* (2010) resulted in changes impacting learning by closing the loop (Banta & Blaich, 2011). Additionally, there are impediments to conducting assessments and then acting upon their findings. Several such challenges include a need for effective, knowledgeable leaders at institutions; the recognition of shared responsibility for learning that includes both learners and teachers; and faculty and staff development. Despite the many challenges, given the similarity of those impacting EBIP development and authentic assessment, addressing them simultaneously may provide an efficient strategy to make progress on both efforts. We share more details about this approach in the section below.

While authentic assessment is a more complex endeavor than implementing national norm-referenced assessments, methods can be applied on campuses that are authentic and provide for comparison to national achievement targets. For example, Puncochar and Klett (2013) detail an inquiry-based approach to assessing liberal education outcomes in STEM general education courses and employ methods to avoid scoring bias when reviewing a randomized selection of samples drawn from natural science courses. While Puncochar and Klett (2013) detail their methods and findings, the study does not include how these results were used to “close the loop” in faculty practice. Other studies, primarily more narrowly focused scholarship of teaching and learning work, provide insight into looping assessment findings back into course design (Grassian, 2013; Kogan & Laursen, 2014; Romeo & Posey, 2013; Rowe et al., 2015). Increasingly, campuses and state university systems are utilizing the authentic assessment and “Valid Assessment of Learning in Undergraduate Education” (VALUE) rubrics (AAC&U, 2014) to support comparable yet authentic assessments of learning outcomes in general education (Pusecker et al., 2011), pushing against entrenched systems and patterns in teaching and assessment.

This study illustrates one institution’s efforts to disrupt established STEM general education teaching practices and to integrate thinking about authentic learning outcomes assessment into teaching. We probed and explored faculty teaching practice through direct observation, asked faculty to report learning outcomes in their courses, and sought connections between those practices. Our results provide insight into the challenges faculty experience in connecting pedagogy, student learning, and assessment, especially in a context in which they are being asked to consider making changes to their teaching practice. They also point to actions we can take to support educational reform.

## Method

### Background

This study took place in the context of broad institutional efforts to increase the adoption of active learning practices and to link those practices to authentic assessment of university learning outcomes (ULOs). It began the year after instituting a new set of Liberal Education and America's Promise–inspired (AAC&U, 2014) ULOs and introducing a newly designed general education program that was integrated into all majors. We used a zero-based curriculum review process to ensure alignment between courses and the outcomes and goals of the new program (Paulsen & Pesea, 1992) and supported backward course design through summer institutes to generate a course design table linking outcomes, practices, and planned assessment strategies (Fink, 2003). The ULOs highlight skills that are natural partners to EBIPs (e.g., communication, problem solving, and teamwork). Additionally, the ULOs we have adopted incorporate five sets of disciplinary outcomes including criteria for mathematics (ULO 7) and the natural, physical, and applied sciences (ULO 8), which are referenced below. Ultimately, the goal for the reform has been to move toward a learning-centered university (Tagg, 2003) characterized by intentional course design, EBIPs, and authentic assessment, which loops back into course improvements (Biggs & Teng, 2011; Fink, 2003).

### Teaching Observation to Foster Disruption and Reflection

In order to understand the connection between pedagogy and student learning outcomes, we were interested in documenting teaching practice directly. We opted for an observation instrument specifically designed for use in STEM classrooms—the Reformed Teaching Observation Protocol (RTOP) instrument (Budd, Van der Hoeven Kraft, McConnell, & Vislova, 2013; Sawada et al., 2002), which highlights evidence-based practices in lesson design, content, and classroom culture. The RTOP fit within the broader shift described above and is noted as “constructivist”; it has proved useful in other STEM reform teaching efforts to prompt discussion about teaching during teaching workshops (Marshall, Smart, Lotter, & Sirbu, 2011). We note the need for observation to be conducted with care and mindfulness of its capacity to be taken as criticism, even among the most confident and experienced faculty (Donnelly, 2007). The RTOP provided fodder for faculty reflection and fostered our capacity to “stir the pot.”

### University Learning Outcomes Assessment

The university's assessment reporting process follows the logic of the course design tables described above, in which each ULO criterion is explicitly linked to pedagogies and assessment methods. The process asks faculty to report assessment data collected, conclusions reached about students' learning from the assessment, and implications those conclusions have for future teaching. The opportunity to probe how faculty understand these connections is a significant driver of this research. Fifteen partial or complete assessment reports were submitted in our study's second year; we did not press participants for outcomes assessment in the third year. Almost all participants submitted some type of grade equivalent score, while far fewer submitted interpretations of those assessments of learning outcomes proficiencies in addition to scores (see Table 1). Open-ended comments by the faculty highlighted the criteria on which they noted students struggling or the need for additional curricular coverage (e.g., “I need better assessment tools,” “I expect to see a higher average as I continue to integrate more active learning in the classroom”).

Table 1.

*Example Yr1 ULO Reporting Data*

*Mathematics course example*

	7.1 Quantitative Reasoning	7.2 Math computation	7.3 Appropriate strategy	7.4 Form, justify generalizations	7.5 Technology as a tool
Grade					
n	<i>n</i> = 8	<i>n</i> = 7	<i>n</i> = 6	<i>n</i> = 7	<i>n</i> = 6
M	<i>M</i> = 56.6	<i>M</i> = 58.9	<i>M</i> = 62.6	<i>M</i> = 63.1	<i>M</i> = 57.4
SD	<i>SD</i> = 34.3	<i>SD</i> = 26.6	<i>SD</i> = 6.7	<i>SD</i> = 27.0	<i>SD</i> = 43.3
Proficiency Score (freq.)					
1	0	1	0	0	0
2	1	1	3	2	2
3	8	4	4	6	2
4	0	0	0	0	4

*Science course example*

	8.1 Process of inquiry	8.2 Use logic, interpretation	8.3 Communicate scientific ideas	8.4 Technology – human interaction
Grade				
n	<i>n</i> = 7	<i>n</i> = 6	<i>n</i> = 6	<i>n</i> = 5
M	<i>M</i> = 61.5	<i>M</i> = 67.2	<i>M</i> = 66.9	<i>M</i> = 70.6
SD	<i>SD</i> = 23.1	<i>SD</i> = 16.7	<i>SD</i> = 14.6	<i>SD</i> = 9.3
Proficiency Score (freq.)				
1	0	1	0	0
2	5	2	6	3
3	3	5	1	3
4	0	0	0	0

Annual Cycles of the Study

We employed an action research model, with annual cycles informing adjustments to the methods for the following year. Funding from a National Science Foundation grant enabled us to offer a small stipend to participating faculty (\$25 gift card) in the first year and \$500 in each of the subsequent two years of the study. The increase was due to modifications to the protocol that required more time and effort from the participants. Second- and third-year participants were also given copies of Angelo and Cross's *Classroom Assessment Techniques* (1993). The study received Institutional Review Board approval for all cycles and modifications of the project. The three years of the study and the modifications made to our procedures are described below.

Each year all faculty teaching the general education STEM courses included in this study were invited to participate. An initial e-mail was sent explaining the study and inviting them to participate; within two weeks research personnel followed up via e-mail and phone calls with personal invitations, often meeting with potential participants to explain the study.

In year 1 (Yr1) faculty were recruited, and trained RTOP observers were assigned to faculty participants to conduct observations. Faculty were contacted by observers via e-mail, and general information on the course was exchanged. RTOP reports were sent to faculty prior to their end-of-semester semistructured interviews. Initial review of the first

few interviews helped us understand the power of the RTOP to potentially destabilize faculty (e.g., “I looked at this and thought . . . , So you just gave me a middle C. . . . I know a lot more about chemistry than that”). We immediately changed our process to contextualize the RTOP results and to provide the results in person. At the conclusion of their interview, faculty were asked to share learning outcomes assessment data in “any form that made sense” to them. Reviewing the formats used and the patterns of their answers helped the research team design the assessment survey used in subsequent years. (See online resources for a description of the reporting survey and instructions to faculty <http://doi.org/10.18122/B24G6Z>.)

In year 2 (Yr2), we added a workshop to introduce participants to the RTOP, to suggest classroom assessment techniques (Angelo & Cross, 1993), and to share authentic strategies for ULO assessment. During the semester, we paired a research team member with each participant for a midterm check-in and then to deliver the RTOP results at the end of the term. Responding to Yr1 participants’ desire for interaction with other study participants, the semistructured interviews were replaced by focus groups designed to provide collaborative conversations simultaneously with debriefing. Data collected included a written worksheet and observation notes taken during the focus groups. Finally, faculty submitted ULO assessment reports using the survey developed in Yr1.

Year 3 (Yr3) followed the Yr2 design, with small adjustments. We shortened the prestudy workshop, and the focus group sessions were recorded and transcribed.

### Participants and Research Site

The participants in this study taught at a large public metropolitan university with a doctoral research designation. As shown in Table 2, three cohorts of faculty over the three years included seventy-one courses in six science and mathematics disciplines of our general education curriculum. It should be noted that the rank of the teaching faculty in the general education STEM courses mirrors that of the university, with approximately 50 percent of the faculty in the adjunctive rank, while this project included only eight (14.5 percent) adjunctive faculty; therefore adjunctive faculty are underrepresented. Mathematics is overrepresented in our sample because, due to its smaller classes, there were many more mathematics instructors who were eligible for participation in the study. In addition, as discussed below, the mathematics faculty had been involved in departmental teaching reform efforts, which also played a role in their willingness to volunteer for the study.

Table 2  
*Distribution of Class Sections and Disciplines*

Discipline	# course sections	Gen Ed STEM		Upper Division
		Science	Mathematics	
Biology	6	5		1
Chemistry	7	6		1
Construction Management	2			2
Engineering Science	5	5		
Geoscience	3	3		
Mathematics	43		40	3
Physics	5	4		1
<b>Total</b>	<b>71</b>	<b>23</b>	<b>40</b>	<b>8</b>

Fifteen participants were involved in multiple years; from this subset of participants we selected illustrative narratives that reflect themes emerging in the research. While the initial data analysis included all fifty-five participants, narrowing the number of cases to multiyear participants for close study enabled us to manage the wealth of data we collected and retain the subjectivities of the participants’ experiences (Tappan, 2001). Pseudonyms have been assigned to participants (see Table 3), and we have grouped the faculty into discipline categories in order to protect their

identities. It should be noted that adjunctive faculty were invited to participate; however, only one (Isabel) participated in more than one year of the study to be included in the current analysis. The remainder of the faculty were full-time lecturers or tenure/tenure-track faculty.

Table 3  
*Participant and Average RTOP Results*

Name	Discipline	Yr1 RTOP	Yr2 RTOP	Yr3 RTOP	# Terms taught	Years participated
Rob	SCI	32	28		64	2
Terry	SCI	19	34		69	2
Ellen	MATH	39	39		18	2
Allen	MATH	36	39		15	2
Eric	MATH	39	32		14	2
Isabel	MATH	37	40		44	2
Ann	MATH	41	43		67	2
Henry	MATH	38		44	43	2
Erin	SCI		25	46	10	2
Gene	SCI	46		56	42	2
Allison	MATH	47		58	66	2
William	MATH	23	59		31	2
Dennis	MATH	41	62		40	2
Trisha	MATH	38	40	66	51	3
Suzanna	MATH	68	48		20	2
Whitney <sup>a</sup>	SCI	39			16	1

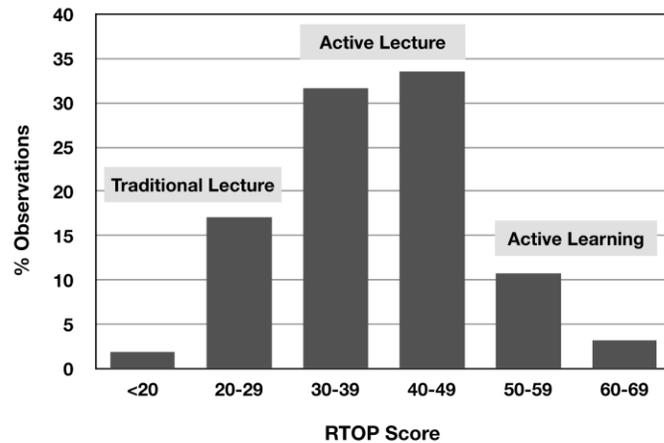
### Data Collection

The twenty-five-item RTOP instrument was administered by trained graduate students following norming sessions with research personnel. Each faculty member was observed by the same graduate student three times during a single semester. We provided every participant a histogram with the average faculty RTOP scores and their own scores (similar to Figure 1). We divided the scores into three descriptive categories:

- Traditional lecture classrooms—RTOP scores less than 30 (instructor-directed lecture with limited student interaction)
- Active lecture classrooms—RTOP scores in the 30–49 range (instructor-directed lecture with some activities involving student-student interactions)

- Active learning classrooms—RTOP scores 50 or greater (significant student-centered activities, especially in small groups, integrated with course content)
- The histogram (Figure 1) helped faculty better understand their RTOP scores in comparison with those of their university peers.

Figure 1  
*Aggregate RTOP Scores*



*Note.* Data represent the distribution of total scores on 25 items for N faculty participants in all years of the study. Participants' scores are averages of all scores obtained in a single semester.

Assessment data included faculty reports of students' proficiency on a scale of 1 to 4 (1 = *unsatisfactory*, 2 = *developing*, 3 = *good*, 4 = *exemplary*) on either the science or the mathematics ULOS, which are aligned with national Valid Assessment of Learning in Undergraduate Education rubrics (AAC&U, 2014). Average grades for assessments connected to the learning outcomes and proficiency scores assigned to student performance were reported (see Table 1). Faculty assessment reports also included three open-ended questions asking faculty to summarize data collected for assessment, conclusions drawn, and implications for their future course planning. Members of the research team conducted initial analyses to determine whether a relationship existed between learning outcomes scores and RTOP scores, and no correlation was detected.

### Data Analysis

Our approach to the data was recursive and ongoing. The process of asking questions of our data in each annual cycle is characteristic of interpretive research; our intention was to highlight and share understandings (Patton, 2002). Our data analysis was aided by the use of NVivo software, which adds transparency to our process (Ryan, 2009) and assisted the ability to make the emerging patterns visible given the sizable amount of data collected. Two of the principal investigators individually identified codes, reviewed data, and refined codes until we reached agreement. Coded data were drafted into memos and shared with our third principal investigator, who probed and challenged us to look at the data anew. Constructed narratives were shared with the larger research team for verification; interim findings were shared in on-campus poster presentations with our study participants and at national teaching conferences (Paterson McGuire & Stieha, 2015; Stieha & Paterson McGuire, 2013). The manuscript of this article was shared with participants for participant checking (Maxwell, 1996) and received their acknowledgment that it represents their experiences. In fact, several participants shared stories about ongoing work refining and polishing the approaches to pedagogies and assessment that they began in this project. Despite the multiple routes that we could have taken through the data, this study explores the ways the study activities prompted participants to think carefully about the pedagogical choices they were making and to explore the connections between authentic learning outcomes assessment and their teaching practices.

## Findings

### The Impact of Observation

Each participant in the study was observed three times during each year of participation, and the impact of being observed emerged as a strong theme in the data. Particularly in Yr1, teaching choices and considerations about the merits or feasibility of active learning practices that were directly tied to the contextual factors surrounding the observed classes filled the transcripts.

Terry was a veteran faculty member teaching a large (>150 students) general education science lecture class. During the Yr1 interview, Terry referred to the RTOP subscores and engaged in an active deliberation on a sense of what “could” or “should” be:

Students’ interaction. I could do better on that. I can imagine in some cases I show them a question and I have them answer it . . . as a class; they’re clicker questions, so I do it as a class. I could have them do it in pairs or triplets and discuss it, and I could probably increase my student-student interaction score, I think in a useful way. But again, having a high proportion of student talk, I don’t think it’s appropriate. I guess I don’t feel like in a class like mine where I’m preparing them to start taking premed classes etc., I feel like there’s certain information that needs to be transmitted and I don’t have enough time for 20–30 percent students to be talking. (Yr1 Interview)

As we listen to this narrative, we are aware that the RTOP had stirred up thinking about active learning: what can be done to increase “student interaction” and a sense of reasonable limits based on the perceived constraints on this particular course (class size, preparation for medical school, expected limited student participation).

In the second year Terry commented, “Being observed bugged me . . . to make an effort to do more active things in the classroom. I did modify my behavior because I was being observed.” Terry used the study as a reason to give active learning a try a few times, writing on the focus group worksheet, “As a result of participation in this study I found myself adding more ‘active activities’ than I have in the past. That is a good result of the study! . . . I hope that my increased classroom ‘activity level’ translates into increased learning.” Terry’s subsequent assessment of student learning reveals the increases she sought: “I can see that I should spend more time on material and activities associated with criteria 8.4. I am pleased with the students’ understanding of the concepts and knowledge of Criteria 8.1 and 8.2” (Yr2 Assessment Report).

Rob, also a veteran science professor, knew about the active learning practices that are being used in the field; he stated that “flipped classes” and “POGIL [Process Oriented Guided Inquiry Learning]” approaches are possible but affirmed reasons not to change his long-held practice of large lectures with pauses for questions. Rob noted the need to cover a large “amount of material,” adding, “I’ve dabbled in new things, in active stuff, but I feel constrained by the amount of material and the number of students, so I settled on a less active way” (Yr1 Interview). He would like, “someday, maybe,” to increase the use of EBIPs.

It is possible that being observed as part of this study left Rob and Terry with a sense that they had to justify maintaining traditional teaching practices, and Terry did give active practices a try. While Terry and Rob were actively deliberating the benefits of EBIPs, they saw limits placed upon them due to the size of their courses and content coverage expectations—elements of departmental or disciplinary cultures.

In the Yr2 focus groups a conversation among participants further highlighted the impact that an observer in the classroom had on their teaching. Two participants wanted the observer to see student-to-student interactions. Three people referenced wanting to give “a clear presentation that day.” A faculty member in mathematics, Suzanna, provided details about what she had in mind: “I was thinking about general organization, communication, and evidence so that the observer knows what the objectives are and that the lesson activities always support that. It was supposed to be a day that we started something new. . . . I wanted to be clear that there was building on prior knowledge” (Yr2 Focus Group). Suzanna added that planning for observation added intentionality to her active learning choices—it “helps you think about what you are doing that day.”

Making teaching public through observation provides the impetus to make teaching choices intentional and to try out EBIPs. Observation, however, did not generally have an impact on faculty thinking about the assessment of student learning. This underscored a general observation from our study: participants did not automatically connect teaching choices to student learning, or they compartmentalized these two activities when discussing them.

### Connecting Teaching Practice to Student Learning

Faculty participants in mathematics were, on the whole, more likely than those in the sciences to see EBIPs as feasible in their classes. Many had smaller classes, and there was a growing departmental culture supportive of active learning in lower-division courses. Allison (a mathematics professor) found building more EBIPs into her practice a welcome challenge. At the same time, she indicated that assessing the ULOs remained an area of uncertainty for her. She mentioned that the mathematics ULO criteria “use appropriate technology as a tool for problem solving” and said, “I don’t have a really good way of assessing; I listen to what they’re saying in class, but I don’t really have a good way of assessing [ULO 7.5]” (Yr1 Interview).

Allison’s comments represent a theme emerging from the Yr1 interviews that even when faculty had the desire to connect practices and assessment, they did not necessarily have the assessment tools or they associated assessment with summative measures. During the interviews we discovered that asking faculty members to connect their work to the plans for teaching and assessment in the course design table (Fink, 2003) for their course did not help them be clearer about the connections between teaching and assessment. These discussions illuminated how those who had not been part of the original course design process were disconnected from the intended learning assessment strategies.

Although her assessment questions lingered, during the Yr3 focus group Allison described the RTOP as a measure to help her gauge improvement in her use of inquiry-based learning: “I specifically tried to do more presentations and target things that are deemed inquiry-based learning. And so, I was pleased that my [RTOP] scores went up for that reason because I was trying to accomplish that, I guess. I don’t know [whether] that means it was a better class or that they learned more, necessarily just from the scores, but it does show that I did accomplish a little bit of what I was trying to do” (Yr3 Focus Group). While Allison’s Yr3 assessment report does not explicitly connect increasing EBIPs to student learning, it does indicate a thoughtful approach to learning outcomes assessment using a blend of formative and summative assessments, including “daily class presentation and discussion of homework” and “problems on the exam.” She concluded from these data, “Students are better at routine problems than non-routine. They need more opportunities to develop ULO 7.2 [communication about mathematical ideas] and 7.3 [appropriate problem-solving strategies]” (Yr3 Assessment Report). Allison referenced needing more formative assessment: “I need to make sure I am giving feedback on written homework relating to ULO 7.2, 7.3, and 7.4 [formulate and justify generalizations]. I could also choose problems to be presented in class that fit these ULOs so that they are sure to be discussed” (Yr3 Assessment Report).

Disconnections between teaching practice and learning outcomes assessment was a persistent theme, yet the results do show movement toward greater understanding of how to connect them. Ann provided a good example of the study activities increasing her confidence in linking assessment and practice over multiple years. A veteran math professor, Ann explained that she had historically focused on course content outcomes that must transfer from Calculus I to Calculus II. She said, “I know what they need in the next course . . . , so that’s always in the forefront of my mind” (Yr1 Interview). Returning to the study in Yr3, Ann focused more on ULO assessment—she said that she was “thinking more about the ULOs this semester” and that doing so helped her “focus on how to ask questions on the final exam” (Yr3 Focus Group). The increased clarity about the connection between EBIPs and assessment is evident: “In choosing to use a larger amount of time with students working in groups on material they had not seen or had just gotten a bit of, I hoped to see students deepen understanding because they had to struggle/explain to others etc.—in looking at my results that I used in my assessment it appears that it did help the students move forward in their understanding” (Yr3 Focus Group Worksheet). Ann’s mention of “material they had not seen” is a direct reference to the mathematics ULO criterion to solve “nonroutine problems and correctly interpret results.” Ann’s experience as a Yr1 and Yr3 participant helps us see the time required for this evolution in her thinking about EBIPs and ULOs.

### Leveraging Faculty-Faculty Interactions

As action researchers, we were alert to the ways our study activities not only were mechanisms for research but fostered conversations among the faculty participants and supported a desired cultural shift on campus. In our data, we see that faculty adoption of the idea of EBIPs as important and achievable preceded shifts in their approach to assessment to demonstrate students' learning progress. Faculty peer conversations in this study helped shift attitudes from resistance to openness to EBIPs and authentic assessment practices. Indeed, this finding echoes that of others in higher education who suggest that the opportunity for faculty to reflect collaboratively is supportive of reform efforts (Cochran-Smith & Lytle, 2009; Rutz et al., 2012; Sellheim & Weddle, 2015).

As we conducted the analysis of these data, relationships with other colleagues, programs, and courses emerged as an important factor for faculty considering EBIPs and assessment. The focus group conversations among peers in Yr2 and Yr3 contributed to many participants' ability to connect EBIPs and student learning. For example, Whitney taught an upper-division course focused on "communication in the discipline," which is a component of the general education curriculum. She commented that the scores on the RTOP were "easily manipulatable" because she "selected the first two [class sessions for observation] with the idea that this was going to be evaluated and I needed to have a lot of student participation." She questioned how "reliable" the scores were and doubted the benefit of active learning. Trisha responded, "You don't see the benefit to the students?" Whitney answered, "Correct. In one instance I did. In others I was doing it for the benefit of [the study] as opposed to . . . the students." Trisha replied, "Last time we did this debrief it came up that there are certain disciplines, maybe overall, or at least certain courses, or maybe parts of courses, where having them work on an activity doesn't suit the material, doesn't help the student, and lecturing kind of is the way to go." Their conversation continued, with the other focus group participants and facilitators listening. Trisha shared that she had learned to try out new approaches. Whitney concluded, "In that respect I think it was good because I definitely tried a couple things that I hadn't necessarily tried before. I think the class needs to have more of those because I thought it was really successful. In that way I think it's [the study] really succeeding" (Yr3 Focus Group).

In the exchange with Whitney, Trisha stepped forward to share her insights and experiences as a former RTOP skeptic who had come to understand that one purpose is stretching pedagogically. Finally, she voiced a theme that was echoed in many narratives throughout the three years of the study: "It's a lot of work, so you kind of need incentive to try it" (Yr3 Focus Group). Trisha's collegial move in the exchange helped to move the conversation from resistance to seeing the possibility in EBIPs.

Interestingly, while we saw the RTOP as valuable to "[give] us a common vocabulary" (Yr3 Focus Group) about EBIPs, participants valued the focus groups. For example, Suzanna clearly favored interaction over the instrument: "It wouldn't be about RTOP for me at all. . . . It is about developing relationships and community and thinking about teaching. . . . I think of it as a way to create opportunities for dialogue" (Yr3 Focus Group). In a related narrative, Erin described her project mentor coming to observe her natural science class and providing feedback to improve student-to-student interactions during Yr2. Erin's written comments clearly point to ideas she gained from the study's presemester workshop and from consulting with an expert in pedagogy: "With [her] feedback, I was able to implement group discussion, individual writing exercises, etc. I noticed that after I asked students to group up the classroom atmosphere changed! Almost instantaneously, more students were responding to my questions (verbally) and classroom participation increased! . . . This was a concept that I learned about during the workshop" (Yr2 Focus Group Worksheet). In terms of planning to collect and look at assessments of student learning data, Erin wrote, "I wrote specific test questions from the . . . lectures designed to be more 'active' and found that students generally (not all) got them correct" (Yr3 Focus Group Worksheet). Reporting on student proficiency, Erin wrote, "Some are getting it . . . and some are not. I have concluded that multiple teaching techniques benefit student learning. For example, some students do well with a traditional lecture style, and do not know how to handle in-class, group exercises, while others benefit from active learning exercises" (Yr3 Assessment Report). Both observation as a catalyst for growth and teaching as the subject of public discourse are benefits of these interactions.

## Discussion

### Connecting EBIPs and Assessment Practices Is a Multiyear Disruptive, Communal Process

One of the benefits of this action research project is that we were able to document the narratives of faculty over a period of time during which they were focusing on change of practice in a variety of classes. The blend of support and accountability over several semesters provided access to those narratives. These conditions are difficult to document if the study is limited to single courses or only follows faculty for a single term (Blaich & Wise, 2015; Henderson et al., 2011). In some of the cases we have discussed above, support, encouragement, and accountability did help faculty integrate EBIPs and link those practices to our ULOs and outcome assessment. In particular, Erin and Ann exemplify how faculty navigate this multistep process. Even entertaining what is pedagogically possible, as in Trisha's, Terry's, and Rob's cases, opened the door to experimenting with new methods and shifting their practices at their own pace.

Further, we were able to discern that the connections between teaching practice and learning assessment practices are made but are very difficult to surface. Ann's, Allison's, and Terry's narratives support the understanding that connecting authentic assessment to teaching practice requires a long view of teaching and learning to be taken. Through detailed analysis of coded narrative data, we were able to draw the line from one year's activities to the next to ultimately see the connection. This work, however, is laborious and still does not result in neat sound bites that are easily digestible by the public or legislative bodies questioning the outcomes of higher education.

### Integrating EBIPs and Assessment Practices Requires Incentives for Faculty

The weight of "time" and "content" in the participants' narratives is clearly a force dragging against more integration of EBIPs and authentic assessment. For example, Trisha's comment underscores the extra workload: learning how to effectively integrate EBIPs into a course is "a lot of work" and requires "incentives" (Yr3 Focus Group). The progression moved from trying an EBIP to gaining confidence and then adjusting assessments to measure desired learning outcomes from the practice. These comments are consistent with prior research (Henderson et al., 2011; Webber, 2012) reminding us of the time and complexity involved in changing teaching practice. Even with incentives, some faculty concluded that there were limits on their ability to integrate EBIPs due to time and content constraints. As a result, faculty chose "less active" EBIPs, such as using real-world examples and "lecture pausing" (Rob), and searched for the "right balance" of active learning (Trisha) given their department culture. In relation to instructor salaries, the incentive of \$500 to participate in the yearlong project was modest, which may indicate that incentives may be more symbolic and tangible.

"Limits" to EBIP use are connected to factors at least partially under a department's control (e.g., class size and content sequencing) when setting standards for its general education courses. This element of department culture remains invisible when it is not surfaced through dialogue, as is evident in Rob's and Terry's narratives. For our mathematics participants, however, the department value on discourse about the balance between EBIPs and content coverage (e.g., Ann's narrative) exposed that tension and invited pedagogical innovation. Our findings thus echo the power of departmental culture cited elsewhere in the literature (Dubrow, 2004; Gess-Newsome et al., 2003; Lund & Stains, 2015) as a dominant factor in pedagogical reform.

### Stirring the Pot Through Social Networks

Finally, comments throughout the study regarding faculty interactions—from the least formal (hallway chats) to those related to intentional project activities (workshops, focus group conversations, discussions with mentors)—suggested that there were social networks within departments and across departments, and associated with the Center for Teaching and Learning, that were supporting the development of innovative teaching and assessment practices. Our Yr3 focus groups, in particular, revealed how interactions, such as the conversation between Trisha and Whitney described above, have the power to shift faculty perspectives. We have been alerted to the reality that change processes emerge and are maintained through interpersonal relationships and find resonance with others who conclude that "lasting change does not result from plans, blueprints, and events" (Mohrman, Tenkasi, & Mohrman, 2003, p. 321). We are now seeking pockets of innovation on campus where we can work to cultivate reform.

## Conclusions and Future Research

This study highlights the importance of incentivizing and supporting faculty to systematically approach the redesign of courses to include evidence-based instructional practices and authentic assessment. It further highlights the importance of departmental cultures that help to determine content (material to be covered and course sequencing), context (class size), and the available time for faculty to engage in reform. Departments also play a role in supporting and aligning ULO assessment activities with internal program assessment. In part due to the activities from this grant-funded research, the university received additional funding from the National Science Foundation's Widening the Implementation and Dissemination of Evidence-Based Reform program to catalyze widespread faculty adoption of EBIPs within STEM departments, where most of our participants noted the greatest prospects to affect change. One important reason to focus efforts at the department level is the potential impact on adjunctive faculty. As noted earlier, adjunct faculty were underrepresented in this project. Because adjunct faculty's network connections are most likely to be in their home department, a more local focus may be the best way to engage adjunct faculty in reform efforts.

As a product of this action research, we have applied our findings in courses across the general education curriculum. We have begun using an observation protocol and mentorship to support faculty to integrate more active learning into their teaching practices and to identify faculty development needs in our first-year seminar courses. We are also expanding reform efforts into art, humanities, and social science courses through a university-funded initiative, "Engaging Pedagogies."

Finally, having seen the power of faculty peer conversations to shift perspectives about EBIPs and assessment, we understand the importance of these relationships in bringing about organic change. As we continue to document change, we hope to increase attention to the highly complicated relationship among the constructs of pedagogy, learning, and assessment. Accordingly, we join others (Daly, 2010; Kezar, 2014; Quardokus & Henderson, 2014) in suggesting that social network theory has great utility in higher education in understanding and fostering cultural change. Our future work includes analysis of social networks on campus that can be leveraged to stimulate reform efforts further.

Vicki Stieha, Ph.D., is an assistant professor of education at Boise State University and serves as the faculty associate for learning outcomes assessment. She earned a Ph.D. in educational foundations at the University of Cincinnati. She conducts research on teaching and learning in higher education, particularly attending to the ways that organizational culture supports and impedes innovation in pedagogy and assessment.

Susan E. Shadle, Ph.D., is the director of the Center for Teaching and Learning and a professor of chemistry and biochemistry at Boise State University. She began her work at the Center for Teaching and Learning in 2006 after ten years as a faculty member in the Department of Chemistry and Biochemistry. Shadle was recognized as the 2015 Carnegie Professor of the Year for Idaho and is the principle investigator on a National Science Foundation Widening Implementation and Demonstration of Evidence-Based Reforms project that provides support to faculty in science, engineering, and math to adopt evidence-based teaching practices. She is active in the national Process Oriented Guided Inquiry Learning community, conducting workshops and leading initiatives for the project. She has co-authored a chemistry text using the Process Oriented Guided Inquiry Learning methods that she uses in her general chemistry course. She is a member of the Professional and Organizational Development Network in Higher Education, where she serves as an ambassador for the local region.

Sharon Paterson is an associate professor of sociology at Boise State. She earned her Ph.D. from Virginia Polytechnic Institute and State University. She was most recently the Vice Provost at Boise State and contributed to record increases in student persistence. Her research interests include investigating inequalities in higher education.

### Note

We gratefully acknowledge the work of Dr. Karen Viskupic for training in the use of the Reformed Teaching Observation Protocol and the collection of data therefrom. We also appreciate the collaboration of the research team (Dr. Eric Landrum, Dr. Sasha Wang, Dr. Doug Bullock, and Ms. Erin Muggli) for their contributions to this research. This material is based upon work supported by the National Science Foundation under grant DUE-#1256585.

## Works Cited

- Andrews, T. C., & Lemons, P. P. (2015). It's personal: Biology instructors prioritize personal evidence over empirical evidence in teaching decisions. *CBE—Life Sciences Education*, *14*(1). doi:10.1187/cbe.14-05-0084.
- Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd ed.). San Francisco: Jossey-Bass.
- Association of American Colleges and Universities. (2014). Liberal Education and America's Promise. Retrieved from <https://www.aacu.org/leap/>.
- Banta, T. W., & Blaich, C. (2011). Closing the assessment loop. *Change*, *43*(1), 22–27.
- Banta, T. W., Jones, E. A., & Black, K. E. (2010). *Designing effective assessment: Principles and profiles of good practice*. San Francisco: John Wiley and Sons.
- Berrett, D. (2012, October 25). Lectures still dominate science and math teaching, sometimes hampering student success. *Chronicle of Higher Education*. Retrieved from <https://crpde.ou.edu/sites/crpde.ou.edu/files/STEM%20Teaching.pdf>.
- Biggs, J. B., & Teng, C. (2011). *Teaching for quality learning at university: What the student does* (4th ed.). Philadelphia: McGraw-Hill.
- Blaich, C., & Wise, K. (2015, January 22). Knowing about vs. knowing how. Center of Inquiry. Retrieved from <http://www.liberalarts.wabash.edu/practitioners-corner/2015/1/22/knowing-about-vs-knowing-how.html>.
- Brownell, S. E., & Tanner, K. D. (2015). Barriers to faculty pedagogical change: Lack of training, time, incentives, and . . . tensions with professional identity? *CBE—Life Sciences Education*, *11*(4), 339–46. doi:10.1187/cbe.12-09-0163.
- Budd, D. A., Van der Hoeven Kraft, K. J., McConnell, D. A., & Vislova, T. (2013). Characterizing teaching in introductory geology courses: Measuring classroom practices. *Journal of Geoscience Education*, *61*(4), 461–75.
- Carmel, J. H., Jessa, Y., & Yeziarski, E. J. (2015). Targeting the development of content knowledge and scientific reasoning: Reforming college-level chemistry for nonscience majors. *Journal of Chemical Education*, *92*(1), 46–51. doi:10.1021/ed500207t.
- Centra, J. A. (1994). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. San Francisco: Jossey-Bass.
- Chang, T.-S., Lin, H.-H., & Song, M.-M. (2011). University faculty members' perceptions of their teaching efficacy. *Innovations in Education and Teaching International*, *48*(1), 49–60.
- Chen, X., & Soldner, M. (2013). *STEM attrition: College students' paths into and out of STEM fields* (NCES 2014-001). Washington, D.C.: National Center for Educational Statistics. Retrieved from <http://nces.ed.gov/pubs2014/2014001rev.pdf>.
- Cochran-Smith, M., & Lytle, S. L. (2009). *Inquiry as stance: Practitioner research for the next generation*. New York: Teachers College Press.
- Daly, A. J. (2010). Mapping the terrain: Social network theory and educational change. In A. J. Daly (Ed.), *Social network theory and educational change* (pp. 1–16). Cambridge: Harvard Educational Press.
- Donnelly, R. (2007). Perceived impact of peer observation of teaching in higher education. *International Journal of Teaching and Learning in Higher Education*, *19*(2), 117–29.
- Dubrow, G. (2004). Collegiality and culture: General education curriculum reform at Western Protestant University. *Journal of General Education*, *53*(2), 107–34.
- Eastwood, J., Sadler, T., Sherwood, R., & Schlegel, W. (2013). Students' participation in an interdisciplinary, socioscientific issues based undergraduate human biology major and their understanding of scientific inquiry. *Research in Science Education*, *43*(3), 1051–78. doi:10.1007/s11165-012-9298-x.
- Eddy, S. L., Coverse, M., & Wenderoth, M. P. (2015). PORTAAL: A classroom observation tool assessing evidence-based teaching practices for active learning in large science, technology, engineering, and mathematics classes. *CBE—Life Sciences Education*, *14*(4). doi:10.1187/cbe.14-06-0095.
- Fink, D. (2003). *Creating significant learning experiences*. San Francisco: Jossey-Bass.
- Fox, M. A., & Hackerman, N. (Eds.). (2003). *Evaluating and improving undergraduate teaching in science, technology, engineering, and mathematics*. Washington, D.C.: National Academy of Sciences.

- 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 of the United States of America*, *111*(23), 8410-8415. doi:10.1073/pnas.1319030111
- Galbraith, C. S., & Merrill, G. B. (2012). Predicting student achievement in university-level business and economics classes: Peer observation of classroom instruction and student ratings of teaching effectiveness. *College Teaching*, *60*(2), 48–55.
- Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education*, *53*, 229–61. doi:10.1007/s11162-011-9247-y.
- Gess-Newsome, J., Southerland, S. A., Johnston, A., & Woodbury, S. (2003). Educational reform, personal practical theories, and dissatisfaction: The anatomy of change in college science teaching. *American Educational Research Journal*, *40*(3), 731–67.
- Graff, G. (2010). Why assessment? *Pedagogy*, *10*(1), 153–65.
- Grassian, D. (2013). Constructive ambiguities: The inspiring, deflating, transformative, and limited possibilities for assessment in higher education. *Journal on Excellence in College Teaching*, *24*(2), 155–73.
- 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), 64–74.
- 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), 952–84.
- Heywood, J. (2000). *Assessment in higher education: Student learning, teaching, programmes, and institutions*. London: Jessica Kingsley.
- Kazempour, M., Amirshokooi, A., & Harwood, W. (2012). Exploring students' perceptions of science and inquiry in a reform-based undergraduate biology course. *Journal of College Science Teaching*, *42*(2), 38–43.
- Kezar, A. (2014). Higher education change and social networks: A review of research. *Journal of Higher Education*, *85*(1), 91–125.
- Kogan, M., & Laursen, S. L. (2014). Assessing long-term effects of inquiry-based learning: A case study from college mathematics. *Innovative Higher Education*, *39*(3), 183–99. doi:10.1007/s10755-013-9269-9.
- Kuh, G. D., Kinzie, J., Buckley, J. A., Bridges, B. K., & Hayek, J. C. (2006). *What matters to student success: A review of the literature*. Retrieved from [http://nces.ed.gov/npec/pdf/kuh\\_team\\_report.pdf](http://nces.ed.gov/npec/pdf/kuh_team_report.pdf).
- Love, B., Hodge, A., Grandgenett, N., & Swift, A. W. (2014). Student learning and perceptions in a flipped linear algebra course. *International Journal of Mathematical Education in Science and Technology*, *45*(3), 317–24. doi:10.1080/0020739X.2013.822582.
- Lund, T. J., & Stains, M. (2015). The importance of context: An exploration of factors influencing the adoption of student-centered teaching among chemistry, biology, and physics faculty. *International Journal of STEM Education*, *2*(13). doi:10.1186/s40594-015-0026-8.
- Marshall, J., Horton, R., & White, C. (2009). EQUIPPing teachers. *Science Teacher*, *76*(4), 46–53.
- Marshall, J., Smart, J., Lotter, C., & Sirbu, C. (2011). Comparative analysis of two inquiry observational protocols: Striving to better understand the quality of teacher-facilitated inquiry-based instruction. *School Science and Mathematics*, *111*(6), 306–15. doi:10.1111/j.1949-8594.2011.00091.x.
- Martinez, J. F., Borko, H., Stecher, B., Luskin, R., & Klower, M. (2012). Measuring classroom assessment practice using instructional artifacts: A validation study of the QAS Notebook. *Educational Assessment*, *17*(2–3), 107–31.
- Maxwell, J. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks: SAGE.
- Mohrman, S. A., Tenkasi, R. V., & Mohrman, A. M., Jr. (2003). The role of networks in fundamental organizational change: A grounded analysis. *Journal of Applied Behavioral Science*, *39*(3), 301–23.
- Murphy, K. L., Picione, J., & Holme, T. A. (2010). Data-driven implementation and adaptation of new teaching methodologies. *Journal of College Science Teaching*, *40*(2), 80–86.
- Olakanmi, E. O., & Doyoyo, M. (2014). Using structured examples and prompting reflective questions to correct misconceptions about thermodynamic concepts. *European Journal of Engineering Education*, *39*(2), 157–87. doi:10.1080/03043797.2013.833177.
- Olson, S., & Riordan, D. G. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. Report to the president*. Retrieved from <http://eric.ed.gov/?id=ED541511>.

- Pakala, K., & Bose, D. (2015). *Use of mobile learning strategies and devices for e-portfolio content creation in engineering thermodynamics and fluid mechanics classes: Student perceptions*. Paper presented at the 2015 American Society for Engineering Education Annual Conference and Exposition, Seattle. Retrieved from <https://peer.asee.org/24978>.
- Paterson McGuire, S., & Stieha, V. (2015). *Teaching observations and faculty development to integrate Evidence-based practices and assessment of student learning outcomes*. Paper presented at the Lilly Conference on Evidence-Based Teaching and Learning, Newport Beach, CA.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Paulsen, M. B., & Pesea, B. A. (1992). A practical guide to zero-based curriculum review. *Innovative Higher Education*, 16(3), 211–21.
- Piburn, M., Sawada, D., Turley, J., Falconer, K., Bloom, I., & Judson, E. (2000). *Reformed teaching observation protocol (RTOP) reference manual*. Retrieved from Tempe, AZ: <http://www.public.asu.edu/~anton1/AssessArticles/Assessments/Biology%20Assessments/RTOP%20Reference%20Manual.pdf>
- Puncochar, J., & Klett, M. (2013). A model for outcomes assessment of undergraduate science knowledge and inquiry processes. *Research and Practice in Assessment*, 8, 42 – 54
- Pusecker, K. L., Torres, M. R., Crawford, I., Levia, D., Lehman, D., & Copic, G. (2011). Increasing the Validity of Outcomes Assessment. *Peer Review*, 13/14(4/1), 27-30.
- Quardokus, K., & Henderson, C. (2014, 1 April). *Using department-level social networks to inform instructional change initiatives*. Paper presented at the National Association for Research in Science Teaching Annual Conference, Pittsburgh.
- Romeo, P., & Posey, F. (2013). Assessment: Can it lead to a better course design? *Journal of College Science Teaching*, 43(1), 30–35.
- Rowe, M. P., Gillespie, B. M., Harris, K. R., Koether, S. D., Shannon, L.-J. Y., & Rose, L. A. (2015). Redesigning a general education science course to promote critical thinking. *CBE—Life Sciences Education*, 14(3). doi:10.1187/cbe.15-02-0032.
- Rutz, C., Condon, W., Iverson, E. R., Manduca, C. A., & Willett, G. (2012). Faculty professional development and student learning: What is the relationship? *Change*, 44(3), 40–47. doi:10.1080/00091383.2012.672915.
- Ryan, M. (2009). Making visible the coding process: Using qualitative data software in a post-structural study. *Issues in Educational Research*, 19(2), 142-161.
- Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R., & Bloom, I. (2002). Measuring reform practices in science and mathematics classrooms: The Reformed Teaching Observation Protocol. *School Science and Mathematics*, 102(6), 245–53.
- Sellheim, D., & Weddle, M. (2015). Using a collaborative course reflection process to enhance faculty curriculum development. *College Teaching*, 63, 52–61.
- Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder: Westview.
- 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 Sciences Education*, 12(4), 618–27.
- Stieha, V., & Paterson McGuire, S. (2013). *Documenting pedagogy and learning in STEM General Education*. Paper presented at the AAC&U, Portland.
- Tagg, J. (2003). *The learning paradigm college*. San Francisco: Jossey-Bass.
- Tappan, M. B. (2001). Interpretive psychology: Stories, circles, and understanding lived experience. In D. L. Tolman & M. Brydon-Miller (Eds.), *From subjects to subjectivities: A handbook of interpretive and participatory methods* (pp. 45–56). New York: New York University Press.
- Watkins, J., & Mazur, E. (2013). Retaining students in science, technology, engineering, and mathematics (STEM) majors. *Journal of College Science Teaching*, 42(5), 36–41.
- Webber, K. L. (2012). The use of learner-centered assessment in US colleges and universities. *Research in Higher Education*, 53, 201–28.
- Wiggins, G. P. (1993). *Assessing student performance: Exploring the purpose and limits of testing*. San Francisco: Jossey-Bass.