6-1-2019

# A Multi-Institutional Analysis of Instructional Beliefs and Practices in Gateway Courses to the Sciences

Joseph J. Ferrare
*University of Kentucky*, joseph.ferrare@uky.edu

# A Multi-Institutional Analysis of Instructional Beliefs and Practices in Gateway Courses to the Sciences

Digital Object Identifier (DOI)
https://doi.org/10.1187/cbe.17-12-0257

Notes/Citation Information
Published in *CBE—Life Sciences Education,* v. 18, no. 2, ar26, p. 1-16.

# A Multi-Institutional Analysis of Instructional Beliefs and Practices in Gateway Courses to the Sciences

**Joseph J. Ferrare***
Department of Educational Policy Studies and Evaluation, Department of Sociology, and Martin School of Public Policy and Administration, University of Kentucky, Lexington, KY 40506-0001

## ABSTRACT

**This paper builds on previous studies of instructional practice in science, technology, engineering, and mathematics courses by reporting findings from a study of the relationship between instructors' beliefs about teaching and learning and their observed classroom practices. Data collection took place across six institutions of higher education and included in-depth interviews with 71 instructors and more than 140 hours of classroom observations using the Teaching Dimensions Observation Protocol. Thematic coding of interviews identified 31 distinct beliefs that instructors held about the ways students best learn introductory concepts and skills in these courses. Cluster analysis of the observation data suggested that their observable practices could be classified into four instructional styles. Further analysis suggested that these instructional styles corresponded to disparate sets of beliefs about student learning. The results add momentum to reform efforts that simultaneously approach instructional change in introductory courses as a dynamic relationship between instructors' subjective beliefs about teaching and learning and their strategies in the classroom.**

## INTRODUCTION

College students who wish to enter academic majors in the sciences must first pass through a notoriously rigorous set of introductory courses. These "gateway" courses are typically taken in the first 2 years of college and may include sequences in general chemistry, physics, and biology, along with courses in mathematics (e.g., calculus, differential equations) and engineering (e.g., computer programming, engineering mechanics). The imagery of "gates" is appropriate, because these courses literally constitute a barrier between incoming students and these selective academic majors and, ultimately, professions. While successfully passing gateway courses does not guarantee degree completion in the sciences, previous research has identified these courses as among the greatest obstacles along this trajectory (Suresh, 2007; Chang *et al.*, 2008; Alexander *et al.*, 2009; Gasiewski *et al.*, 2012; Malcom and Feder, 2016).

Early investigations of persistence in the sciences suggested that instructional practices in introductory courses were a primary culprit in driving students away from these majors (Gainen, 1995; Seymour and Hewitt, 1997). More recently, researchers have found that the adoption of instructional practices that actively engage students significantly bolsters student performance in science, technology, engineering, and mathematics (STEM) courses (Freeman *et al.*, 2014). This research is part of a broader movement among policy makers and educational reformers to leverage instructional practices as a way to foster a more diverse and talented supply of scientists (Wieman *et al.*, 2010; President's Council of Advisors on Science and Technology, 2012). The primary assumption guiding these reform efforts is that the traditional "sage on a stage" approach of imparting the foundational concepts and skills of science is driving many students to other areas of study who would otherwise persist in STEM fields.

**Marilyne Stains,** *Monitoring Editor*
Submitted Dec 8, 2017; Revised Dec 13, 2018; Accepted Feb 19, 2019
CBE Life Sci Educ June 1, 2019 18:ar26
DOI:10.1187/cbe.17-12-0257
*Address correspondence to: Joseph J. Ferrare (joseph.ferrare@uky.edu).

The importance of instructional practices in reforming STEM education has led to an extensive research agenda focusing on multiple facets of classroom instruction (Henderson *et al.*, 2011). A central component of this agenda has involved a push to descriptively catalogue practices in STEM courses (Hora, 2015). These efforts have led to the development of multiple classroom observation instruments (for a review, see Hora and Ferrare, 2012), such as the Reformed Teaching Observation Protocol (i.e., RTOP; Sawada *et al.*, 2002), Practical Observation Rubric to Assess Active Learning (Eddy *et al.*, 2015), Teaching Dimensions Observation Protocol (Hora and Ferrare, 2014b), and the Classroom Observation Protocol for Undergraduate STEM (i.e., COPUS; Smith *et al.*, 2013)—and in some cases combinations of these protocols (e.g., see Lund *et al.*, 2015).

Studies using these and other protocols (including self-report instruments) have generally challenged attempts to dichotomize instructional practices into traditional lecturing versus active engagement, and instead have found that teaching practices in STEM courses often contain multiple dimensions of practice and forms of engagement (Gasiewski *et al.*, 2012; Hora and Ferrare, 2014a). Given the diverse set of fields that make up "STEM," a substantial portion of this work has occurred within discipline-based education research communities. Physics education researchers, for example, have documented a wide range of instructional practices in introductory courses, whether focusing on single institutions (West *et al.*, 2013) or across a broad sample of institutions (Dancy and Henderson, 2010). Chemistry education researchers have also found substantial variety in the types of instructional practices used in chemistry courses, including small-group work, interactive styles, and multiple modes of lecturing (Gibbons *et al.*, 2018). In the context of the geosciences, meanwhile, the range of practices identified by the RTOP did not correlate with any instructor demographics or institutional characteristics (Teasdale *et al.*, 2017).

Researchers have also documented instructional practices across STEM disciplines and have generally come to similar conclusions regarding the variability of practices and forms of engagement (Smith *et al.*, 2014; Swap and Walter, 2015; Drinkwater *et al.*, 2017). The most comprehensive examination of instructional practice in undergraduate STEM courses was conducted by Lund *et al.* (2015), who used the COPUS and RTOP instruments to code 269 class periods from a sample of 73 instructors across 28 research-intensive universities in the United States. The study by Lund and colleagues found that the class periods clustered into 10 COPUS profiles that ranged from different lecturing formats (e.g., didactic, Socratic) to collaborative learning situations (e.g., peer interaction, group work). More than half (52%) of observed class periods were classified as lecture-centric, and that number increased to more than two-thirds once Socratic lecture styles (i.e., Q&A with students) were included.

In addition to describing instructional practices found in STEM courses, researchers have sought to examine instructors' beliefs about teaching and learning in these contexts (Feldman, 2000; Harwood *et al.*, 2006; Lotter *et al.*, 2007; Henderson and Dancy, 2008). Some of the earliest work on this topic was conducted by Prosser *et al.* (1994), who examined instructors' beliefs about teaching and student learning in the context of introductory physics and chemistry courses. Their study found that instructors tended to espouse beliefs of student learning that ranged from conceptual development and change to objective knowledge acquisition. Subsequently, the instructors' beliefs about teaching tended to follow a similar underlying schema. More recently, Hora (2014) found that STEM instructors' beliefs could be arranged along a teacher-centered versus student-centered continuum, but that most instructors in the sample espoused components from across these dimensions. For example, while most instructors believed that "practice and perseverance" was a crucial student-centered component of learning, many who held this belief also espoused instructor-centered beliefs such as scaffolding or providing clear explanations of content.

The body of work focusing on beliefs about teaching and learning has tended to conceptualize subjective interpretations not as ancillary components to practice, but rather as fundamental components of pedagogy. Indeed, Woodbury and Gess-Newsome's (2002) widely recognized teacher-centered systemic reform (TCSR) model places instructor thinking at the center of influence when conceptualizing instructional change (see also Gess-Newsome *et al.*, 2003). In particular, the TCSR model assumes that instructors' beliefs about teaching, learning, and content are inextricably linked to their classroom practices.

Researchers seeking to explore this connection within and between STEM fields have found that the relationship between instructional practice and beliefs about teaching and learning emerge through a complex set of cultural assumptions and institutional settings (Sunal *et al.*, 2001; Hora and Hunter, 2014; Marbach-Ad *et al.*, 2014). For instance, Gibbons and colleagues (2018) used surveys to examine the connection between instructional practices and beliefs across a large sample of chemistry faculty. These researchers found that instructors who facilitated interactive and small-group work styles of instruction held the strongest student-centered beliefs about learning relative to instructors with more lecture-centric styles. Meanwhile, those whose practice was characterized as lecturing with clicker response systems tended to fall between these groups with respect to their beliefs about learning.

Other researchers have examined instructional practices and beliefs across STEM fields. Using an interview-based approach, Ferrare and Hora (2014) found that STEM instructors' practices in the classroom emerged through interactions between their beliefs about how students learn and the constraints encountered within classroom, departmental, and disciplinary environments. That is, instructors appear to have tacit theories of teaching and learning that are enacted imperfectly due to their perceptions about what can be accomplished in practical settings of the university (see also Hora, 2014; Lund and Stains, 2015; Stains *et al.*, 2015; Stains and Vickrey, 2017). To understand the complexities of instructional practice, then, prior research suggests it is important to link instructors' subjective beliefs about teaching and learning to the practices they use within the constraints of their academic milieus (Kane *et al.*, 2002).

While researchers are accumulating insightful evidence about instructional practices in STEM courses, less work has connected these practices to instructors' beliefs about teaching and learning within the specific context of introductory (or "gateway") courses in these fields. Although Ferrare and Hora (2014; see also Hora, 2014) did explore this link in STEM courses, their study focused primarily on belief systems and

only explored the connection to practice among a subsample of three faculty. Others have taken a more systematic approach by focusing on this link within disciplines (e.g., respectively, in physics and chemistry, see Dancy and Henderson, 2010; Gibbons *et al.*, 2018). The following paper works to build on this area of the literature by examining beliefs about teaching and learning alongside instructional practices among 71 instructors of introductory STEM courses across six institutions of higher education in the United States. In particular, the analysis addressed the following questions:

1. What are instructors' beliefs about how students best learn foundational concepts, processes, and skills in introductory STEM courses?
2. What observable practices do instructors in these courses use to facilitate student learning and engagement?
3. How do instructors' beliefs about teaching and learning in introductory courses relate to their observed practices?

These questions were designed to facilitate an in-depth analysis of instructional beliefs and practices in introductory courses that serve as gateways to STEM majors. Rather than testing specific hypotheses about the extent to which instructional practices conformed to a certain continuum or set of categories (e.g., interactive vs. lecture based), this study sought to document emergent patterns of beliefs and practices with as few assumptions as possible. In addition, the combination of interviews and observations allowed for an exploration of the ways instructors' beliefs and practices interact "in the wild." In the process, the present study helps inform current reform efforts by adding another layer to the descriptive account of instructional practices in introductory STEM courses across several types of institutions (research, private, liberal arts, etc.).

## METHODS

The present study used a multi-institutional case study design (Yin, 2008), drawing on data collected between 2012 and 2014 from introductory STEM courses at six institutions of higher education (IHEs) across multiple geographic regions of the United States (Mountain West, Midwest, and East Coast). Data collection consisted of more than 140 hours of classroom observations in 71 introductory courses, and semistructured interviews with the 71 instructors of record for each course (see Table 1 for sample characteristics).[1] The IHEs varied in mission, size, and selectivity, consisting of three flagship research universities (>30,000 students), one non-flagship research university (>30,000), a medium-sized (<15,000) private university, and a small (<5000) private liberal arts college. While the selection of institutions is by no means representative of all IHEs in the United States, the maximum variation sample used in the study offered the capacity to examine introductory courses across a wide variety of organizational and geographic settings.

The sampling unit for this study consisted of introductory courses that serve as gateways to STEM majors. Courses most likely to be gateway courses were initially chosen based on a review of the literature (e.g., Seymour and Hewitt, 1997; Suresh, 2007; Alexander *et al.*, 2009; Gasiewski *et al.*, 2012; Malcom and Feder, 2016) and by examining course requirements for entry into STEM majors at the participating sites. At each site, these courses typically included: general biology, general physics (calculus and algebra based), general chemistry, organic chemistry, calculus 1–3, differential equations, introduction to programming, and data structures. This initial list was then circulated to academic advisors, instructors, and other informants at each institution to ensure that all introductory courses were included and to add those that might be unique to a particular site.

The data collection was carried out by a team of four researchers, all of whom had advanced degrees in education research at the time of the fieldwork.[2] Site visits to each institution lasted 2 weeks and typically took place near the middle of the term so as to observe practices once the instructors and students had a chance to establish some degree of routine. During the 2-week visit, each instructor was interviewed once and the course was observed twice.[3] The instructors were informed about the exact dates and times of the observations. The observer sat in the back of the room in an attempt to

**TABLE 1. Instructor attributes in the sample of introductory courses**

| | *N* (%) |
|---|---|
| Sex | |
| Male | 47 (66) |
| Female | 24 (34) |
| Racial/ethnic identity | |
| White | 55 (78) |
| Asian or Pacific Islander | 5 (7) |
| Latino/a or Hispanic | 2 (3) |
| Black or African American | 0 (0) |
| Native American or Alaska Native | 1 (1) |
| Not reported | 8 (1) |
| Discipline group | |
| Biology | 9 (13) |
| Chemistry | 18 (25) |
| Computer science | 7 (10) |
| Engineering | 10 (14) |
| Mathematics | 14 (20) |
| Physics | 13 (18) |
| Job title | |
| Teaching assistant | 2 (3) |
| Lecturer or instructor | 26 (37) |
| Senior lecturer or senior instructor | 5 (7) |
| Visiting professor | 2 (3) |
| Assistant professor | 6 (8) |
| Associate professor | 16 (23) |
| Professor | 13 (18) |
| Other | 1 (1) |

[1]The data collection for this project was part of a larger data-collection effort that consisted of one-on-one and focus group interviews with students. The results from the interviews and focus groups are forthcoming elsewhere.

[2]The research team is identified in the *Acknowledgments*. Due to the extensive scope of the project and dissemination efforts, the team democratically decided to limit authorship to those who contributed to data analysis and writing of the article (see the Vancouver Protocol: http://storage.googleapis.com/wzukusers/user-17415557/documents/56640b2c61339C4KMzWo/Vancouver%20Protocol.pdf).

[3]Nine of the courses met once per week for 3 hours. In these cases, the course was only observed once for the full 3 hours.

minimize his or her presence and potential impact on the instructor and students. The interviews were usually conducted between the first and second observation, although scheduling conflicts dictated that the interview fell outside this sequence in some cases. Sequencing the interviews and observations in this way may have prompted the instructors to think more critically and prepare differently for their practice, which may in turn have altered their behaviors in the classroom. However, the advantage of this approach was that it allowed an interviewer to have a frame of reference for instructors' responses and ask more specific questions about past and future actions related to their instructional practices and beliefs about student learning in the context of the course.

The instructor of record for each course was contacted via email solicitation and upon consent was scheduled for a 90-minute interview. The interviews covered a wide range of topics, including beliefs about teaching and learning in introductory courses, factors that influence teaching practices, and broader views about persistence in STEM fields. The present analysis focused on the segment of the interview that investigated instructors' views about teaching and learning in introductory courses. This portion of the interviews was semistructured and prompted by the following questions:

- What are the most important things you want students to learn in [the course]?
- Is there anything about the nature of [most important thing(s) mentioned] that suggests a specific approach or style of teaching?
- What do you think is the best approach to introducing students to [most important thing(s)] in this course? What role, if any, does the instructor play in this approach?
- What is your view about how undergraduates come to understand and apply the [most important thing(s)] in this course?

Although these questions provided a general structure to the interviews, the interviewers were trained to investigate emergent themes through follow-up questions.

To collect classroom observation data, this study used the Teaching Dimensions Observation Protocol (TDOP; Hora and Ferrare, 2014b). The TDOP is similar to the COPUS protocol (Smith *et al.*, 2013), in that it captures instructionally relevant activities at 2-minute intervals. Unlike the RTOP (Sawada *et al.*, 2002), which aims to measure the use of specific reform practices, the TDOP (and COPUS) captures a wider variety of activities related to instructional practices (e.g., lecturing with premade visuals, small-group work), pedagogical moves (e.g., use of humor, adding emphasis), interactions (e.g., display questions, peer interactions), cognitive engagements (e.g., problem solving, creating), and technology use (e.g., PowerPoint slides, digital tablet).[4]

The range of activities for which the analyst is responsible for coding when using the TDOP also creates a significant burden in terms of interrater reliability (Smith *et al.*, 2013).[5] To address this limitation, the four members of the research team engaged in a multiday training that included a thorough review of the codes and extensive observation of videotaped introductory classes in biology, chemistry, physics, and mathematics. The videotaped classes were publicly available via YouTube and were selected based on the quality of video (clarity, scope of camera angles, etc.) and breadth of instructional practices represented both within and between class periods. The selection criteria were meant to expose the raters to the widest possible variety of practices that the TDOP is meant to capture.

At the conclusion of the training, the team coded four additional videotaped classes—two in general chemistry and two in calculus—and achieved an average Cohen's kappa of 0.70 across all pairs of raters and dimensions of the TDOP (i.e., instructional practices, pedagogical moves, interactions, cognitive engagements, and technology). However, there was substantial variation in agreement across different dimensions of the instrument. Similar to previous studies using the TDOP (Hora and Ferrare, 2013; Hora, 2015), the raters achieved high levels of Cohen's kappa when coding instructional practices (0.90) and technology use (0.85) and lower levels when coding pedagogical moves (0.56), interactions (0.63), and cognitive engagements (0.56). As a result, only a selected set of codes (i.e., more reliable) from the TDOP were used in the analysis of observation data (see details below in the section *Clustering of Observation Data*).

### Interview Coding Analysis

Transcripts of the interviews were imported into NVivo software for coding. The analysis began with an open coding of these data to identify recurring concepts (Glaser and Strauss, 1967). Two coding analysts—consisting of the author and a paid graduate assistant—simultaneously worked through a sample of randomly selected transcripts to develop an initial set of 41 concept codes (Saldana, 2013). Following multiple revisions to the codebook to ensure consistent specificity of the concepts, the two analysts separately applied the concept codes to another random set of transcripts from each site and revised ambiguous concept codes through discussion (MacQueen *et al.*, 2008). This process continued until the two analysts consistently reached a minimum 70% match rate using a Jaccard similarity measure. The latter match rate indicates the proportion of instances in which both analysts applied a code to the same text when at least one of the two applied a code (Gower, 1985). The two analysts then applied the codebook to the entire data set following the principles of the constant comparative method (Corbin and Strauss, 2008). Once this process was completed, the author reviewed all text fragments for each code as a secondary procedure to ensure consistency in the application of the codebook. In the end, the final codebook consisted of 34 concept codes related to instructors' beliefs about the most important content and skills students should learn in their courses and how they believe students best learn such content and skills.

### Clustering of Observation Data

The primary objective of the analysis of observation data was to inductively classify the sample of courses into mutually exclusive groups distinguished by the frequency with which certain combinations of practices (i.e., TDOP codes) were observed. To

[4]See http://tdop.wceruw.org for more information, including a copy of the instrument.

[5]The COPUS may facilitate higher rates of interrater agreement, because it does not include as many fine-grained distinctions or cognitively based assessments as the TDOP.

accomplish this task, the analysis built upon previous cluster-analytic approaches to analyzing instructional practices in higher education (Stes and Van Petegem, 2014; Halpin and Kieffer, 2015; Lund *et al.*, 2015). Cluster analysis refers to a set of techniques that attempt to iteratively classify objects (e.g., variables or cases) into mutually exclusive groups based on a measure of (dis)similarity between each pair of objects.

The cluster analysis of observation data followed multiple steps. First, for each of the 71 courses, a row profile was created that illustrated the percentage of total class time (i.e., across both observed class periods) spent on each TDOP code. This approach differs from Lund *et al.* (2015), who created profiles for every individual class period, which resulted in multiple profiles per instructor. The latter approach has the advantage of treating each class period as a distinct event. However, because the present study sought to explore the connection between the instructors' beliefs and practices, it was more appropriate to treat the instructor as the primary unit of analysis.

While the unit of analysis differed from Lund *et al.* (2015), the cluster analysis of TDOP profiles followed a similar process. As in this earlier study, our analysis proceeded by selecting the TDOP codes to be used as clustering variables. There is no consensus about the required ratio of variables to cases in cluster analysis, but some researchers have referred to Formann's (1984) suggestion of using a minimum of $2^k$ cases, where $k$ refers to the number of variables. Lund *et al.* (2015) used this as a criterion for their analysis, which resulted in the selection of 8 COPUS codes. Applying the same rule of thumb to a sample size of $N = 71$ suggested that six TDOP codes should be selected for the present analysis.

Reducing the number of TDOP codes followed a variety of strategies. First, codes falling within the less reliable categories (i.e., pedagogical moves, interactions, and cognitive engagements) were excluded in order to strengthen the reliability of the analysis. However, steps were taken to include proxies of such categories when possible. For instance, retaining instructional practice codes such as "interactive lecture" and "small-group work" offered reliably measured proxies for sustained interactions between instructors and students that were less reliably measured in the interactions dimension of the TDOP. Next, codes that were considered redundant were also excluded. For example, technology codes such as PowerPoint and chalkboard were removed, because they were strongly correlated with the instructional practices of lecturing with premade visuals (0.863, $p < 0.000$) and handmade visuals (0.695, $p < 0.000$), respectively. Finally, rarely used technologies such as overhead projectors, movies, and simulations were removed, while technologies such as clickers and digital tablets were retained. Clickers also offered another reliably measured proxy for student engagement. In the end, the six codes used for the cluster analysis included: lecturing with premade visuals (LPV), lecturing with handmade visuals (LHV), small-group work (SGW), interactive lecture (LINT), clickers (CL), and digital tablet (DT). Table 2 provides the proportion of all 2-minute intervals in which each TDOP code was observed, including the six selected for the cluster analysis.

Hierarchical cluster analysis using average linkage (Sokal and Michener, 1958) was then used to partition the cases into mutually exclusive groups using squared Euclidean distances as the proximity measure. This agglomerative procedure begins with each course as an independent cluster and proceeds iteratively until all courses are grouped as a single cluster. The average linkage method is generally considered to be robust to potential outliers and different cluster structures (Everitt *et al.*, 2011). As with all clustering methods, average linkage does not include a statistical test indicating the number of clusters that best fit the data. However, the dendrogram in Figure 1 allows for

**TABLE 2. Percentage of two-minute intervals each TDOP code was observed across the sample of introductory courses ($N = 71$)[a]**

| | % | SD |
|---|---|---|
| Teaching methods | | |
| Lecture | 13.0 | 14.8 |
| **Lecture: premade visuals** | **37.2** | **36.1** |
| **Lecture: handmade visuals** | **56.9** | **33.3** |
| Lecture: demonstration | 4.3 | 9.8 |
| **Lecture: interactive** | **4.9** | **12.7** |
| **Small-group work** | **13.2** | **21.2** |
| Desk work | 6.5 | 10.6 |
| Class discussion | 0.1 | 0.4 |
| Multimedia | 1.1 | 3.9 |
| Student presentation | 0.8 | 2.9 |
| Pedagogical moves | | |
| Movement | 11.6 | 20.8 |
| Humor | 10.1 | 10.3 |
| Reading | 0.3 | 2.0 |
| Illustration | 18.4 | 21.5 |
| Organization | 4.0 | 5.1 |
| Emphasis | 6.2 | 10.3 |
| Assessment | 9.3 | 13.5 |
| Administrative task | 6.0 | 4.4 |
| Instructor/student interaction | | |
| Rhetorical question | 8.8 | 9.6 |
| Display question | 44.1 | 21.9 |
| Comprehension question | 13.3 | 10.7 |
| Student question | 22.5 | 16.0 |
| Student response | 41.4 | 21.8 |
| Peer interaction | 14.6 | 21.0 |
| Cognitive engagement | | |
| Retain/recall information | 36.7 | 22.7 |
| Problem solving | 34.7 | 25.3 |
| Creating | 3.4 | 13.3 |
| Connections | 25.8 | 24.4 |
| Instructional technology | | |
| Poster | 0.4 | 2.3 |
| Book | 0.5 | 2.1 |
| Notes | 9.0 | 19.5 |
| Pointer | 9.7 | 21.1 |
| Chalk/whiteboard | 47.4 | 38.4 |
| Overhead projector | 1.5 | 6.1 |
| PowerPoint/slides | 33.6 | 37.2 |
| **Clickers** | **5.7** | **11.4** |
| Demonstration equipment | 3.8 | 9.3 |
| **Digital tablet** | **14.4** | **29.3** |
| Movie | 1.3 | 4.8 |
| Simulation | 0.9 | 3.9 |
| Web | 0.9 | 5.8 |

[a]Codes in bold font represent the six TDOP codes selected for cluster analysis.

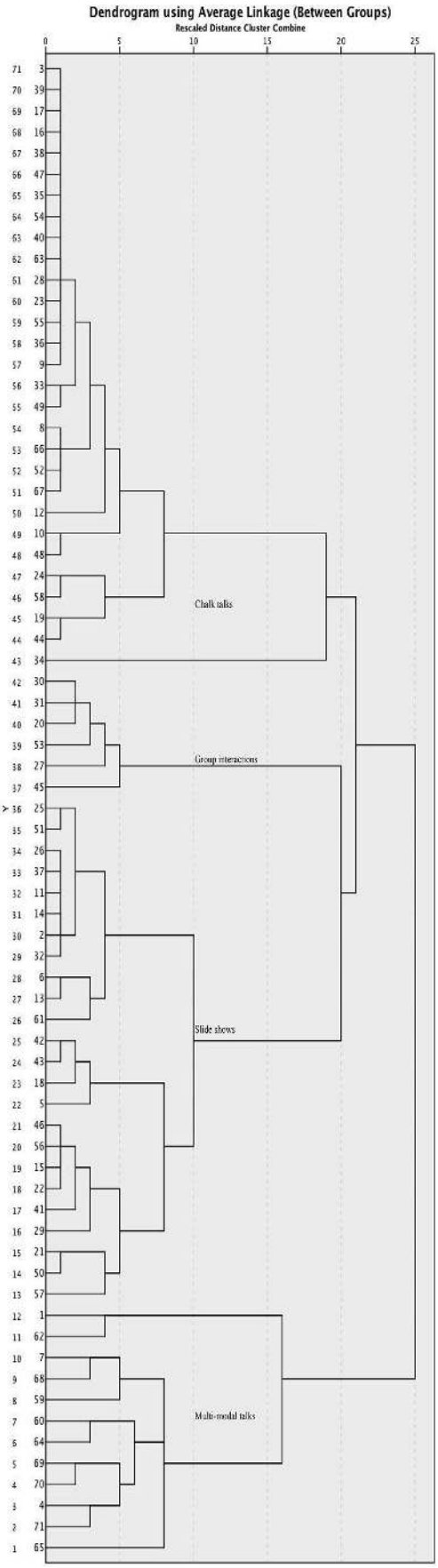


FIGURE 1. Dendrogram of average linkage (between groups) clustering of $N = 71$ courses based on six TDOP codes: LPV, LHV, LINT, SGW, CL, DT (see text for definitions).

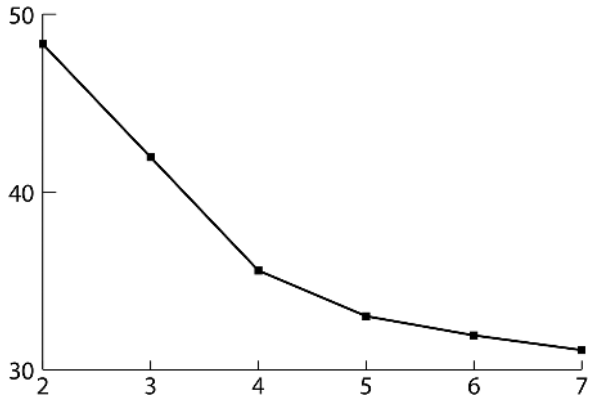


FIGURE 2. Scree plot of change in average distance from cluster centers using *k*-means cluster analysis of $N = 71$ courses based on six TDOP codes: LPV, LHV, LINT, SGW, CL, DT.

an evaluation of the distances between clusters across the stages of agglomeration. In this case, it can be seen that a relatively large increase in distance occurs after the objects (i.e., courses) are clustered into four distinct partitions. As discussed later, these four clusters were selected to represent the types of instructional practices observed in the 71 courses in the sample.

There is no statistical test for selecting the number of clusters, but it is possible to use other clustering routines to test for robustness. A secondary approach used here was *k*-means (MacQueen, 1967), which was selected due to the its prior use in this area of the literature (Lund *et al.*, 2015). *k*-means is a nonhierarchical method of partitioning objects into a distinct set of clusters based on the nearest cluster mean (MacQueen, 1967). This approach does not produce a dendrogram, but it is possible to test multiple numbers of clusters and observe the change in the average distance from the cluster centroids. The scree plot in Figure 2 illustrates this change, ranging from two to seven clusters. Similar to the average linkage approach, the *k*-means clustering suggests that the "payoff" of extending beyond four clusters is minimal. While this does not confirm that a four-cluster solution was the best possible fit for the data, it is suggestive that this solution was not merely an artifact of the specific clustering method used.

### Exploring Interview Codes within and between Observation Clusters

The final stage of the analysis involved an exploration of the relationship between instructors' beliefs about teaching and learning and their observed practices in the classroom. This step involved an analysis of the qualitative codes in relation to the observation clusters and proceeded through three steps. First, matrix coding was used to assess the distribution of each concept code within and between the clusters. The objective of this step was to examine the codes that tended to be overrepresented, underrepresented, and equally represented among the different clusters relative to the overall (i.e., expected) distribution. All of the concept codes were examined during this process regardless of their frequency.

Next, cluster analysis was used to explore the degree of similarity between the concept codes and clusters of observational practices (cf. Ferrare and Hora, 2014; Hora, 2014). Hierarchical clustering with average linkage was again used (Sokal and Michener, 1958), but Jaccard similarity served as the proximity measure, because these data are binary (i.e., presence or absence of a code or cluster; Gower, 1985). The dendrogram was then examined to identify the distinct clusters of concept codes and observational clusters that tended to co-occur in the Jaccard similarity proximity matrix.

The first two steps provided a general overview of the combinations of beliefs about teaching and learning that were associated with instructors' observed practices. As a final step, the raw interview transcripts were partitioned into the same groups as the observational (i.e., TDOP) clusters. Using the coding stripes feature in NVivo, it was then possible to reread the transcripts with an eye toward the ways instructors made connections between their beliefs about learning and the practices they used in the classroom. Attention was given to both the explicit connections that instructors made between their beliefs and practices and the more implicit references that illustrated the ways that their beliefs and practices were intertwined.

## RESULTS

The results are presented in three sections. First, the results from the instructor interviews are reported concerning their beliefs about the ways students best learn introductory concepts and skills in introductory courses. Next, the classroom observation data are presented with special attention to the four clusters that defined the instructional practices within the sample. The final section brings these two results together by illustrating how beliefs about teaching and learning varied within and between the different practice clusters.

### Instructors' Beliefs about Teaching and Learning in Introductory Courses for STEM Majors

During the interviews, instructors were asked about the most important things that students should learn in their respective introductory courses. The interviewers used the generic term "things" so as to allow the widest possible responses. The coded responses to this question initially fell into two broad categories (see Table 3): content acquisition versus skill acquisition. The majority (67.6%) of instructors in the sample pointed to one or more content-oriented concepts within the disciplinary context of a course. The concepts were generally specific, such as series and sequences (calculus), stoichiometry (chemistry), or harmonic oscillations (physics). In most instances, the instructors simply listed key concepts covered in a course, but in some cases they situated the content within a trajectory of understanding, "So if we can leave them with the idea that just because thermodynamics means it should happen, doesn't necessarily mean it will" (senior lecturer, general chemistry).

While learning specific concepts was seen as an important objective in introductory courses, instructors juxtaposed content acquisition with a range of skills that were often seen as more important than the actual content itself. The most frequently cited skill was coded as "conceptual understanding and application" (47.9%). In this case, instructors felt strongly that students needed to be able to discern the underlying concept of a phenomenon to solve unfamiliar problems. For example, a calculus lecturer noted, "We focus less on just memorizing and using formulas and more on understanding the ideas behind calculus … We want you to show the thinking process, the logical process." This sentiment was shared widely across content areas and was seen as a defining feature of college-level work relative to the forms of rote memorization that were perceived to typify high school course work.

A related skill was identified as "perseverance in problem solving" (23.9%), which spoke to the need for students to overcome repeated failure in tackling scientific problems. On many occasions, instructors used the imagery of pushing forward until "the light bulb goes on":

> When reading this text, read with pencil in hand, draw figures to help your understanding, after reading through an example, close the text, try to reproduce the example, if you cannot reproduce it identify where you went wrong, study the text, try again. Stop only when you can comfortably solve the example problem. They [students] have never read a book like that, in that way. And that, I hope they learn by the time they finish this course. That they learn to keep at it until the light bulb goes on, and it will, they're not stupid, for the most part.—Senior lecturer, engineering

This belief in perseverance was frequently connected to conceptual understanding and application and the perception that these qualities were lacking in high school contexts.

**TABLE 3. Coded responses concerning instructors' beliefs and assumptions about the most important things students should learn in gateway courses**

| Codes | %[a] | Description |
|---|---|---|
| Content reference | 67.6 | The instructor referenced content specific to the discipline (e.g., series and sequences). |
| Conceptual understanding and application | 47.9 | Students should learn the underlying concepts (theoretical knowledge) and the different types of contexts in which the content is applicable and know how to identify when such application is prudent so they can apply the concepts to solving problems they have never seen before. |
| Perseverance in solving problems | 23.9 | Students need to learn how to solve problems. In particular, they need to learn how to dig in and grind through tough problems when the answer seems difficult or unobtainable. |
| The identity of "doing science" | 16.9 | Students need to learn how to be a scientist, which is a collaborative process that involves feedback, interaction, deliberation, etc. |
| Connections to daily experience | 12.7 | Students need to learn that the concepts from the course can be experienced in the activities constituting their everyday lives. "Science is everywhere." |
| Interpretation | 5.6 | Students need to learn how to engage with data and tell the story. |

[a]The percentages reflect the number of instructors rather than coded references.

**TABLE 4. Coded responses concerning instructors' beliefs and assumptions of how students best learn key concepts and skills in gateway courses**

| Codes | %[a] | Description |
|---|---|---|
| **Things students do** | | |
| Practice | 39.4 | In order for students to learn the key concepts, processes, and skills from the course they need to practice solving problems in a wide variety of scenarios and contexts. |
| Conceptual application | 33.8 | Learning occurs when students come to understand the underlying concepts and apply these concepts and processes to a wide variety of contexts and problem scenarios and/or draw from existing knowledge and apply it to new problems that have not yet been encountered. |
| Individual perseverance | 29.6 | Students learn when they encounter difficulty and intellectual adversity on their own and have to "grind away" at problems before coming to understand the key underlying principles. |
| Resourcefulness | 16.9 | Students need to learn how to make use of the resources they have available, such as office hours, help desks, teaching assistant, online tutorials, etc. There is no reason students should not do well given the amount of resources available for them to succeed. |
| Connections | 15.5 | Students learn when they connect course material and processes to other courses and everyday situations. |
| Collaboration | 14.1 | Students learn best when they collectively work to solve problems. |
| Explanation & discussion | 12.7 | Students come to understand important concepts and processes when they explain in words what is happening rather than simply providing a formula or solution to a problem. This can include students actively discussing ideas and problems with other students and the instructor. |
| Intellectual risk-taking | 9.9 | Learning involves taking risks by asking questions, engaging, participating, and being willing to get things wrong. This happens in a variety of contexts, such as group work and labs, whole-class scenarios, etc. |
| Apprenticeship | 5.6 | Students learn through acquisition, in which they start with basic skills then proceed to journeyman and ultimately go off to solve their own problems (i.e., mastery) |
| **Things instructors do** | | |
| Provide problem scenarios | 38.0 | Learning is best facilitated when instructors provide opportunities for students to actively solve problems through classroom activities and coherent and challenging assignments. |
| Motivate relevance | 33.8 | Learning is facilitated when the instructor promotes the relevance of concepts and processes and presents them as interesting (i.e., taps into students' curiosity). |
| Demonstrate and model | 25.4 | One of the best ways to introduce students to the most important concepts and processes is by providing in-class demonstrations that the students can experience. This also involves demonstrating the different applications for which the concepts and processes are relevant. |
| Scaffolding | 22.5 | An effective way to introduce students to key concepts and processes is by connecting the material to other concepts and processes they have previously encountered. Sometimes this involves ideas from previous courses, while in other instances it involves building from basic ideas to complex ones. |
| Examples | 21.1 | Learning is facilitated when the instructor provides many examples of the concept or process. |
| Variability | 16.9 | Students learn in a variety of different ways, and there is no single, ideal pedagogical practice. Thus, the best way to introduce students to foundational concepts and processes is to expose them to many different ideas and through many different practices. |
| Theory to application | 15.5 | Learning is best facilitated when the instructor introduces the general theoretical concept and then moves on to apply the theory to solve a variety of problems. |
| Establish rapport and accessibility | 14.1 | Students need an instructor who is approachable so that they feel comfortable asking questions. Being approachable in this context involves an element of instructor fallibility so that students are not intimidated to take a risk by asking questions. |
| Socratic dialogue | 9.9 | Learning is best facilitated through questions posed by the instructor. |
| Repetition | 8.5 | Students need to be introduced to important concepts and processes through repeated exposure. |
| Clear explanations | 8.5 | Learning is best facilitated when ideas and processes are clearly explained with carefully chosen words that connect to students' thinking patterns and experiences. |
| Analogies | 5.6 | Learning is best facilitated when instructors provide analogies between course content and things we encounter in our everyday lives (e.g., negative pressure in the lungs is like pulling a bicycle pump). |

[a]The percentages reflect the number of instructors rather than coded references.

The juxtaposition of content acquisition and skill acquisition also structured instructors' explanations about how students best acquire the most important content and skills, especially through their beliefs concerning the differential role played by instructors and students in the learning process. The top half of Table 4 provides a list of the concept codes reflecting beliefs in which instructors placed the onus of control for learning on students. The concept of "practice" (39.4%) was the most common belief about how students best learn the key concepts and skills in introductory courses. "I tell them it's like sports," a calculus instructor stated, "If you want to be a good swimmer, you have to swim laps. It's the same way in math."

The belief in practice as a way to learn introductory concepts and skills was often coupled with beliefs that learning occurs through individual perseverance (29.6%) during the process of conceptual application (33.8%) in solving problems—the very

same skills that many instructors stated as the most important things to learn in their courses. That is, students were believed to learn through a specific type of practice that involved some degree of struggle over identifying the conceptual structure of a problem and applying it in the correct context. For instance, a lecturer in physics described how he always provided students with all the formulas so their attention would be directed to the underlying concepts. He believed that "your job as a student is [to figure out] 'Is this an energy problem?' 'Am I asking about energy?' [Am I] 'asking about force?' 'What do I have to use to think this through?'" For many instructors, this type of conceptual application was only achieved when persevering through struggle. Referring to his own experience, a computer science instructor reminisced, "Where I learned the most was not where I went in and got the answer immediately, but where I had to struggle to get the answer."

Although less common than practice, individual perseverance, and conceptual application, the need for students to be resourceful was an important belief of many instructors (16.9%). These instructors pointed to a wide range of existing resources that students already have available to succeed, such as help desks, office hours, and online tutorials. As a physics lecturer explained, "Given the other resources that the university has, supports in terms of help desk time and then my office hours and TAs, that somebody does not get an A or a B in this class is an indication that they have not put out enough effort."

Some instructors (14.1%) also pointed to collaboration with other students as a key foundation to learning introductory concepts and skills in STEM courses. "Science is necessarily a collaborative process these days," a biology instructor claimed. "Nobody does science by themselves, and I think both in terms of learning the content … [and] providing that support network that they have peers that they can turn to and help each other learn the material … can be accomplished by students working in teams" (teaching associate professor, biology). Similarly, some instructors (12.7%) believed that explanation and discussion between students and instructors helped to facilitate deep understanding of the key concepts and acquisition of the most important skills. Within this general social environment, others specified beliefs about the importance of making connections across concepts (15.5%), taking intellectual risks (9.9%), and undertaking an apprenticeship (5.6%) as a means to facilitate learning in introductory courses.

While instructors expressed beliefs about what students needed to do to facilitate learning, they also emphasized many beliefs that, at least in part, placed the responsibility of learning on the instructor (see bottom half of Table 4). The most commonly cited strategy was to "provide problem scenarios" (38.0%) in which instructors facilitate skill development through thoughtfully crafted classroom activities and assignments that students experience as challenging and rewarding. In reference to designing clicker questions, a lecturer in chemistry described how "sometimes the teaching is more like 'You need to figure it out through the clicker questions' instead of 'I'm gonna tell you directly.' And so yeah, … I do that on purpose a lot of times." For many, this was seen as a broader strategy that connected to their beliefs that students needed to practice and apply concepts in the process of problem solving. "You teach them something and then you give them a problem that's not exactly the same," one instructor described, "and whether they can solve that problem using the concepts … It's the most telling way that … they really understand" (associate professor, chemistry).

In addition to providing problem scenarios, promoting the relevance of the content was seen as critical to learning by one-third (33.8%) of the instructors in the sample: "Instead of saying, 'We're going to talk about voltage,' which they're still confused about three weeks in, saying, 'This is the application that I want to talk about. This is why you care.'" The task of motivating students to see the relevance of the content was often seen as an initial step in a series of strategies that included scaffolding of concepts, demonstrating phenomena in class (25.4%), and providing a variety of examples (21.1%) for the students to experience. Similar to providing examples, the need for repetition, clear explanations, and analogies to facilitate understanding were viewed as important by a smaller number of instructors: "When you talk about the negative pressure breathing in our own lungs, what is that like? Well, it's like pulling in a bicycle pump" (associate professor, physiology).

While instructors often emphasized beliefs about presenting the content, some also expressed beliefs about facilitating learning through instructor–student interactions. For example, 14.1% believed that instructors had to seem approachable so that students would feel comfortable asking questions and taking intellectual risks. While this may not directly relate to learning, a physics lecturer noted, "If you have some kind of a relationship with the students, then all of these other things [scaffolding, providing problems] are easier to do." The need for Socratic dialogue (9.9%) was another way in which instructors expressed this belief in interaction. However, the latter belief was more of a direct pedagogical technique than a strategy of building rapport with students—although rapport was sometimes seen as a benefit of such dialogue. In reference to learning proofs, a calculus lecturer explained how "the questions are asked by the instructor until he manages the students to … have them wonder about a contradiction, and then induced conclusion from this."

### Instructional Practices in Introductory Courses for STEM Majors

The preceding section provided a general explication of the beliefs that a sample of instructors of introductory STEM courses held about the ways students learn in these contexts. In this section, the presentation of results focuses on the observable practices that these instructors implemented in the classroom. Table 5 and Figure 3 illustrate the distribution of the six TDOP codes across the four practice clusters identified through the cluster analysis. The most common type of instructional practice found in the sample of introductory courses was the "chalk talk," which represented 40.8% ($n$ = 29) of all courses. As the name implies, these courses were characterized by extensive use of lecturing while writing on a chalkboard or whiteboard (81.0% of observed 2-minute intervals), while the use of slides was almost never present (4.1%). Modern technology use in these classrooms was limited overall, with the use of clickers and digital tablets being observed in only 0.4 and 2.9% of 2-minute intervals, respectively. Students in chalk talks rarely interacted with one another through small-group work (4.4%). However, chalk talks represented the highest frequency of interactive lecture (8.5%) in which the instructor facilitated an extended and additive session of Q&A with the students. Chalk talks were overrepresented in math courses (37.9%) and

**TABLE 5. Average proportion of 2-minute intervals in which each TDOP code was observed within each of the four instructional styles**

| | | TDOP code[a] | | | | | |
|---|---|---|---|---|---|---|---|
| **Instructional style (*N*/%)** | | **LPV** | **LHV** | **LINT** | **SGW** | **CL** | **DT** |
| Chalk talks (29/40.8) | Ave. % | 4.1 | 81.0 | 8.5 | 4.4 | 0.4 | 2.9 |
| | SD | 7.7 | 19.4 | 18.6 | 8.5 | 1.7 | 11.6 |
| Slide shows (24/33.8) | Ave. % | 69.3 | 32.5 | 3.0 | 11.1 | 10.8 | 2.3 |
| | SD | 21.7 | 25.7 | 5.4 | 13.3 | 14.7 | 6.8 |
| Multimodal talks (12/16.9) | Ave. % | 63.8 | 71.9 | 0.9 | 10.7 | 8.9 | 73.0 |
| | SD | 29.9 | 18.8 | 2.0 | 13.7 | 13.1 | 22.9 |
| Group interactions (6/8.5) | Ave. % | 15.2 | 8.8 | 3.2 | 69.5 | 4.5 | 1.3 |
| | SD | 12.5 | 9.7 | 6.3 | 19.1 | 11.0 | 2.2 |
| Total (71/100) | Ave. % | 37.2 | 56.9 | 4.9 | 13.2 | 5.7 | 14.4 |
| | SD | 36.1 | 33.3 | 12.7 | 21.2 | 11.4 | 29.3 |

[a]See text for definitions.

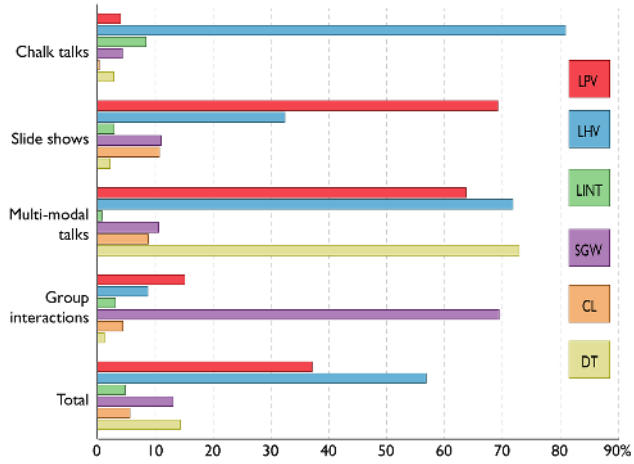


**FIGURE 3. Bar graph of the average proportion of 2-minute intervals in which each TDOP code was observed within each of the four instructional styles.**

underrepresented in biology courses (0.0%), $\chi^2$ (5, $N$ = 71) = 15.73, $p < 0.05$, and were used in expected proportions across the other disciplines in the sample (see Table 6). In addition, chalk talks were equally distributed across all class sizes, $\chi^2$ (2, $N$ = 71) = 1.18, $p > 0.05$, ranging from small (25 or fewer students, 27.6%), medium (26–99, 34.5%), and large (100+, 37.9%).

"Slide shows" were the next most common form of course observed, representing one-third (33.8%, $n$ = 24 courses) of the courses in the sample. While writing at the board was not uncommon (32.5% of 2-minute intervals), instructors in these courses spent most of the time (69.3%) presenting material from premade PowerPoint slides. In addition, instructors facilitating slide shows more frequently included real-time assessments through the use of clickers (10.8%). Students in these courses spent more than twice as much time engaged in small-group work (11.1%) than did their peers in chalk talks. Furthermore, in direct contrast to chalk talks, slide shows were overrepresented in biology (20.8%) and physics (29.2%) and underrepresented in math (0.0%), $\chi^2$ (5, $N$ = 71) = 11.80, $p < 0.05$. Similar to chalk talks, though, there was a consistent distribution of class sizes within the slide show cluster, $\chi^2$ (2, $N$ = 71) = 0.20, $p > 0.05$.

The third cluster—multimodal talks (16.9%, $n$ = 12 courses)—represented a strong overlap between chalk talks and slide shows. Indeed, instructors in these courses tended to vary their mode of delivery between premade and handmade visuals in relatively similar proportions (63.8 and 71.9% of 2-minute intervals, respectively). However, the defining feature among these instructors was the use of the digital tablet (73.0%) as the medium through which the handmade visuals were presented. In addition, students attending multimodal talks interacted through small-group work (10.7%) and answered questions through the use of clickers at a relatively high frequency (8.9%). Although multimodal talks made up

**TABLE 6. Instructional practice clusters by course discipline and class size**

| | **Group interactions** | **Slide shows** | **Chalk talks** | **Multimodal** | **Total** |
|---|---|---|---|---|---|
| Discipline | $N$ = 6 | $N$ = 24 | $N$ = 29 | $N$ = 12 | $N$ = 71 |
| Chemistry | 33.3% | 29.2% | 20.7% | 25.0% | 25.4% |
| Math | 0.0% | 0.0% | 37.9% | 25.0% | 19.7% |
| Physics | 16.7% | 29.2% | 13.8% | 8.3% | 18.3% |
| Biology | 33.3% | 20.8% | 0.0% | 16.7% | 12.7% |
| Computer science | 0.0% | 8.3% | 10.3% | 8.3% | 8.5% |
| Engineering | 16.7% | 12.5% | 17.2% | 16.7% | 15.5% |
| Total | 100% | 100% | 100% | 100% | 100% |
| Class size | | | | | |
| <25 | 33.3% | 25.0% | 27.6% | 0.0% | 22.5% |
| 26–99 | 16.7% | 33.3% | 34.5% | 33.3% | 32.4% |
| 100+ | 50.0% | 41.7% | 37.9% | 66.7% | 45.1% |
| Total | 100% | 100% | 100% | 100% | 100% |

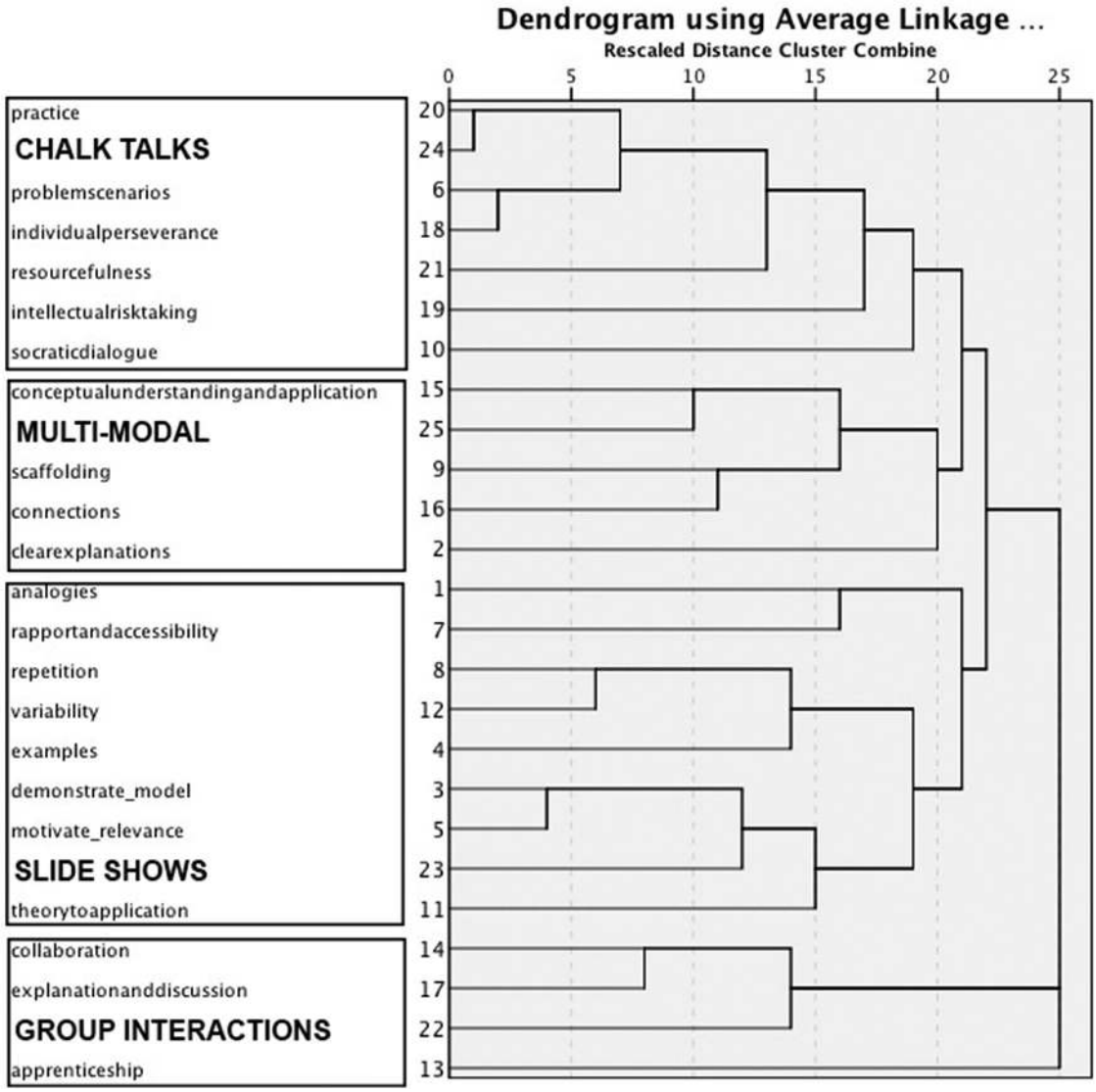


FIGURE 4. Dendrogram of average linkage (between groups) clustering of instructional practice clusters and belief concept codes.

only 12 of the courses in the sample, they were distributed in expected proportions across the disciplines. However, none of the multimodal talk courses were observed in small classrooms (i.e., < 25 students), but were instead concentrated in medium-sized (i.e., 26–99, 33.3%) and larger classrooms of more than 100 students (66.7%), $\chi^2$ (2, $N = 71$) = 4.75, $p < 0.10$.

In strong contrast to the previously mentioned clusters, instructors in courses defined by "group interactions" (8.5%, $n = 6$ courses) rarely used any form of lecture, whether it was with premade slides (15.1%), handmade visuals (8.8%), or sustained interactive lecture (3.2%). Instead, as the name suggests, students in these courses experienced frequent peer interaction through small-group work (69.5%). In these courses, instructors spent considerable time moving from group to group and engaging directly with students. This was in direct contrast to all other clusters, in which the boundary between instructor and student space was clearly defined and consistently maintained. Group interactions were observed among courses in chemistry ($n = 2$), biology ($n = 2$), physics ($n = 1$), and engineering ($n = 1$), and across all sizes of classrooms ($n = 2$ small, $n = 1$ medium, and $n = 3$ large).

### The Intersections of Instructional Practices and Beliefs about Teaching and Learning

Thus far, the reporting of results has focused on instructors' beliefs and practices separately. This is informative for gaining a general understanding of how instructors of introductory courses think about student learning and the strategies they use to facilitate this learning in the classroom. However, it was argued at the outset that these two components of instruction are inextricably linked and thus must also be considered relationally (Woodbury and Gess-Newsome, 2002; Gess-Newsome *et al.*, 2003).

Figure 4 illustrates the dendrogram from the hierarchical clustering of the concept codes and practice clusters. Each of the four practice clusters (i.e., chalk talks, slide shows, group

interactions, and multimodal talks) is positioned around a unique set of concept codes that reflect instructors' beliefs about teaching and learning. As can be seen, instructors who facilitated chalk talks in their classrooms placed the greatest emphasis on beliefs related to practice and providing problem scenarios to facilitate such practice through individual perseverance. As part of this process, these instructors also emphasized the importance of intellectual risk-taking and the use of Socratic dialogue—a practice that was frequently observed in chalk talk classrooms. These beliefs were often expressed as being part of the same process whereby students took intellectual risks by posing and responding to questions through dialogue.

In essence, the instructors' practices during the chalk talk coincided with what they believed students should be doing to facilitate their own learning. That is, the idea of a scientist assiduously working alone to solve a problem with nothing but their thoughts and an empty chalkboard was an ideal that served as both a model of instructional practice and a theory of learning. As one chalk talk instructor described, "Eventually you have to go off and do your own … project … [we try to] get them to the point where they can tackle their own projects and their own problems without having to have someone else tell them where to look for their resources to find those solutions" (lecturer, computer science). For "chalk talkers," the instructor's role in this process was to model problem solving through examples and demonstrations during class to facilitate students' own practice of developing conceptual understanding of the key concepts and skills in the field.

Slide show instructors, meanwhile, believed their role was to promote the relevance of content and subsequently demonstrate and model the content to facilitate student understanding. As part of this process, many slide show instructors believed in the importance of introducing theoretical concepts before discussing any specific application or example. As a lecturer in computer science noted, "What I will do is first go through the theory and the mathematics of it and then go through the application, and then have them take it a step further, basically solving the same type of problem." Taken together, the co-occurrence of slide show instructors' beliefs and practices reflect a latent model of instruction in which the role of the instructor is to theoretically frame the content and then to model examples of the concept through repetition and variability.

Although appearing in a different cluster, slide show instructors also emphasized the presentation of problem scenarios to facilitate conceptual understanding and application. However, these instructors relied more heavily on clicker questions and subsequent student discussion as a way to facilitate this process. To be sure, the instructor remained central in these classrooms, as evidenced by the extensive use of lecturing from PowerPoint slides, and like many others they held a strong belief in the importance of students' individual perseverance. Yet, rather than working through multiple examples at the board as way to model the practice of problem solving, these instructors showed a greater propensity to approach that task through technology (i.e., clickers).

A distinct feature of the multimodal instructors' beliefs was the emphasis on conceptual understanding and application. Whereas other instructors articulated beliefs that conceptual understanding is developed through problem scenarios, multimodal instructors believed more strongly in scaffolding and making connections with other concepts. There remained a strong belief in the importance of practice, but there was a perception that a precursor to students practicing content acquisition involved a variety of pedagogical strategies on the part of the instructor. In particular, the need to demonstrate content or provide problem scenarios was perceived to be less important than figuring out ways for students to connect to the material in a way that resonated with their internal motivations. These instructors often pointed out that, because students' motivations and interests in the content vary widely, it is necessary to present the material through a similarly diverse range of practices. This was clearly seen in the observation of these instructors' classroom practices, which traversed between multiple modes of delivery (i.e., lecturing through slides, handwritten material on the tablet, and small-group work).

As is evident in Figure 4, instructors who facilitated group interaction courses stood apart from all others for their beliefs that collaboration and discussion are fundamental components of student learning. Indeed, these beliefs matched what students were most often observed doing in the classroom. The central theme connecting these beliefs and practices was rooted in social interactions between the instructor and students and directly between students. The action of students explaining and discussing concepts to other students, for example, was perceived as integral to acquiring deep understanding of content and skills. This deeply held belief was directly translated into classroom practices (e.g., small-group work) that facilitated such action. By contrast, none of these instructors expressed a belief in the need to promote the content or practice in the way that was so prevalent among instructors in the other courses. Providing students with problem scenarios was seen as crucial, but rather than facilitating problems through the instructor or technology, these instructors believed in and relied upon student collaboration to carry out this work.

## DISCUSSION

The primary objective of this study was to describe a sample of instructors' beliefs about teaching and learning alongside the observable practices they used in introductory classroom spaces. The results extended prior research in this area in two ways. First, the findings added to a growing body of work that catalogues instructional practices in introductory courses in STEM (e.g., Gasiewski *et al.*, 2012; Hora and Ferrare, 2013; Hora, 2015; Smith *et al.*, 2014; Lund *et al.*, 2015; Swap and Walter, 2015; Drinkwater *et al.*, 2017). In the absence of a representative sample of all IHEs, it is important to observe instructional practices in these courses across several types of institutions (research, private, liberal arts, etc.). Second, the findings connected these observable practices to instructors' underlying beliefs about student learning and the role of both the instructor and students in that process. Understanding these beliefs can inform instructional reform efforts, because such beliefs constitute a type of practical sense among the instructors who are ostensibly the key levers to such transformation (Woodbury and Gess-Newsome, 2002; Gess-Newsome *et al.*, 2003; Ferrare and Hora, 2014; Gibbons *et al.*, 2018).

The results from the analysis of classroom observation data extended the work of Lund *et al.* (2015), who found that instructional practices in a sample of STEM courses at

research-intensive universities could be characterized into four broad instructional styles: lecturing, Socratic, peer instruction, and collaborative learning. Within these styles, their data clustered into a total of 10 unique COPUS profiles (e.g., lecture with slides, limited peer instruction, group work). The findings from the present study using TDOP data both overlapped with and departed from Lund *et al.*'s (2015) work. For example, 8.5% of the courses in the present sample were classified as "group interactions." Similarly, 8.7% of the courses in Lund *et al.* (2015) fell into the "collaborative learning" instructional style that included group work and student-centered peer-instruction COPUS profiles. While the present study used a different instrument and sampling strategy, the similarity of findings in this regard is nevertheless noteworthy.

While 8.5% of courses were classified as group interactions, the vast majority were either chalk talks (40.8%) or slide shows (33.8%). That is, three-quarters (74.6%) of the observed courses were characterized by extensive lecturing at the board or the use of PowerPoint slides, respectively. In this sense, there was relatively limited variability in the types of practices observed in introductory courses across six colleges and universities in the United States. In Lund *et al.*'s (2015) study, 68.4% of the observed class periods were characterized by similar forms of lecturing. In both studies, the use of slides or chalk was not the only difference among these styles of lecturing, though. Chalk talks, and to a lesser extent slide shows, made use of interactive lecturing techniques, just as a substantial number of courses in the Lund *et al.* (2015) study were characterized as Socratic.

Thus, while the overwhelming majority of class periods in both samples fit a limited range of clusters, it would be a mistake to conclude that these STEM courses were either lecture based or interactive, passive or active, or any other simplistic dichotomy (cf. Hora and Ferrare, 2014a; Smith *et al.*, 2014). Instead, the growing body of literature drawing on STEM classroom observations suggests that common lecturing styles vary dramatically in the ways they incorporate student engagement (e.g., Q&A, peer discussion), technology (e.g., clickers, tablets, slides), and cognitive engagements (e.g., problem solving, memorizing; Hora and Ferrare, 2013; Hora, 2015; Smith *et al.*, 2014; Lund *et al.*, 2015; Swap and Walter, 2015; Drinkwater *et al.*, 2017). While prior studies have sought to compare student outcomes in courses with lecture-based versus interactive-based practices (Freeman *et al.*, 2014), the variance found within lecture-based classrooms suggests that future research should also examine whether or not some types of lecturing are more effective than others.

Beyond expanding the understanding of observable classroom practices in introductory STEM courses, the present study contributed to the literature by examining instructors' beliefs about how students learn and the role of the instructor in the process. Many of the beliefs identified in the present analysis were also found in prior studies of STEM instructors. Most notably, Hora (2014; see also Ferrare and Hora, 2014) found that "practice and perseverance" was the most common belief about student learning among a sample of 56 STEM faculty spread across three research-intensive universities. Similarly, the present study identified "practice" and "perseverance" as distinct yet highly prevalent beliefs held by instructors of introductory STEM courses. Other less common beliefs—such as scaffolding, examples, and making connections—were also co-identified across these studies. The prevalence of beliefs concerning practice and perseverance suggests a pervasive belief system among STEM faculty that conceptualizes student learning as a labor-intensive process of "grinding away" at conceptual problems until mastery is achieved. Supporting the pervasive beliefs in practice and perseverance was a more varied set of beliefs about how instructors can facilitate student learning (e.g., scaffolding, application, collaboration), but none were nearly as frequently cited across studies as practice and perseverance.

Prior research has also established that these beliefs are a fundamental component of practice and efforts to reform instructional strategies (Gess-Newsome *et al.*, 2003; Woodbury and Gess-Newsome, 2002). The results from the present study offered a comprehensive look at these beliefs and the intersection with observable practices within the context of a maximum variation sample of introductory courses across STEM disciplines. Similar to prior research (Prosser *et al.*, 1994), instructors' beliefs in this context tended to align toward student-centric or instructor-centric practices that promote learning (see Table 4). However, the cluster analysis of beliefs and practices (see Figure 4) illustrated that some instructors espoused beliefs about student learning that cut across both ends of the spectrum. For instance, instructors who facilitated multimodal talks in their classrooms often held student-centric beliefs related to making connections and conceptual understanding, while also holding an instructor-centric view about the importance of providing students with clear explanations and scaffolding of content. This finding is consistent with previous work examining variability in faculty beliefs about teaching and learning across STEM fields (Ferrare and Hora, 2014; Hora, 2014) and within specific STEM disciplines (e.g., Gibbons *et al.*, 2018).

The results also demonstrated that the instructional styles observed in the classroom tended to correspond to a distinct and coherent set of beliefs about teaching and learning. Previous case studies taking an in-depth look at individual instructors established that instructors' beliefs play an important role in shaping their classroom practices (Ferrare and Hora, 2014; Hora, 2014). Disciplinary-based examinations also found that instructional practice clusters tend to correspond to at least some distinct beliefs about learning and pedagogical self-efficacy (Gibbons *et al.*, 2018). The present study extended prior work through a systematic analysis across a broad sample of STEM faculty teaching introductory courses. The analysis revealed that instructors who practiced the two most common instructional styles—chalk talks and slide shows—expressed disparate beliefs about student learning despite adopting lecture-centric approaches to teaching. On the one hand, chalk talk instructors' beliefs positioned the instructors as the facilitators of student practice through the working out of problems at the board and subsequently posing problem scenarios to students. With slide shows, the instructors perceived themselves to be the facilitators of knowledge by motivating relevance and conveying key concepts through starting with theory and working toward application. Instructors of group interactions and multimodal courses also embodied unique beliefs about teaching and learning that logically coincided with their classroom practices. Differences between instructors, therefore, not only reflected a divergence of classroom behaviors, but also beliefs about how students best understand foundational knowledge in STEM fields.

Understanding the belief systems that inform instructional practices in this sample of introductory courses has implications for reform efforts in these contexts (Harwood *et al.*, 2006; Lotter *et al.*, 2007; Lund and Stains, 2015). Indeed, the push to transform how instructors introduce students to foundational concepts in STEM is not simply a technical problem of changing instructional methods. In addition, such a task necessarily involves an appeal to the ways in which instructors conceptualize the learning process (Wieman *et al.*, 2010)—regardless of whether that understanding has empirical merit. Note, though, that appealing to one's practical understanding of a given practice is not the same as validating that understanding. Rather, it involves building a bridge between practical sense-making processes within a community of practice and the theory of action underlying a desired change (Woodbury and Gess-Newsome, 2002). In this sense, it is impossible to craft meaningful interventions that challenge cultural norms without appealing to the pre-existing meaning systems that are, by definition, already persuasive to the actors involved (Holland and Quinn, 1987). The findings presented here offer a point of departure for such efforts.

### Limitations

This study has several limitations that should be considered when interpreting the results. First, although the sample was designed to maximize variation across a range of institutions of higher education, the results cannot be generalized to the population of instructors of introductory STEM courses in the United States. Instead, these findings should be considered a piece of a broader effort to catalogue instructional beliefs and practices. The present study extended this effort by focusing exclusively on introductory courses in both research-intensive and liberal arts settings. Future work should seek to include additional types of institutions, such as regional or comprehensive universities and community and technical colleges. The latter types of institutions serve a crucial role in educating an expanding and diverse student population, and research in these contexts can help deepen the effort to understand practices and beliefs in STEM fields.

Second, although the use of cluster analysis was an appropriate tool for the present study, it is important to reiterate that this technique does include a statistical hypothesis test to evaluate the fit of the data. It is possible that the four clusters chosen to classify courses in the sample are not the best possible solution. The sensitivity analyses used in the study do suggest that the four clusters were not simply an artifact of the average linkage method. However, future research seeking to test statistical hypotheses about the underlying theoretical constructs that drive instructional practice should attempt to use methods that allow for a direct testing of fit (e.g., latent profile analysis; see Campbell *et al.*, 2017).

Finally, instructors' beliefs were characterized through a qualitative coding of data collected through one-on-one interviews. As a result, there is the potential for bias to emerge during the data collection and analysis. For instance, interpersonal dynamics between the interviewer and interviewee can lead to responses that follow social desirability bias. In addition, qualitative coding involves researchers' subjective interpretations that inevitably include assumptions and biases. These forms of bias were addressed by using semistructured interview protocols to ensure each interviewee was initially prompted by the same questions, but it is still possible that follow-up questions proceeded in different directions depending on the interviewers' own interests and perceptions. The bias associated with data analysis was addressed by using multiple coders and repeated checks to verify consistency in the application of the codebook. The use of surveys can help address some of the shortcomings of interview-based research (e.g., see Dancy and Henderson, 2010; Gibbons *et al.*, 2018), but such methods, of course, have their own limitations.

## CONCLUSION

The findings from this study made use of a maximum variation sample to expand upon what was previously known about instructional beliefs and practices in introductory STEM courses. The findings related to the classroom observation data suggested that instructional practices in the sample of introductory STEM courses could be classified into a relatively few number of instructional styles (i.e., chalk talks, slide shows, multimodal talks, and group interactions). Following prior work, these instructional styles generally varied between student-centered and instructor-centered practices (Lund *et al.*, 2015). The vast majority of the courses in the sample aligned most closely to the latter end of the spectrum by relying heavily on instructor-centered delivery and relatively little direct student-based group work or collaboration.

Instructors' beliefs about teaching and learning also tended to fall along an instructor-centered and student-centered spectrum, although not in a mutually exclusive way (Ferrare and Hora, 2014; Hora, 2014). In the process, this study deepened the literature by focusing on the connection between observable practices and subjective beliefs within the context of introductory courses that students are likely to encounter when pursuing a wide variety of STEM degree programs. This connection adds further support to prior claims that reform efforts must expand beyond the emphasis on technical strategies of instruction to also include the set of beliefs instructors draw upon to inform their practices and how they interpret and subsequently shape instructional reforms (Coburn, 2001; Spillane *et al.*, 2002; Wieman *et al.*, 2010; Lund and Stains, 2015; Stains and Vickrey, 2017).

## ACKNOWLEDGMENTS

This research was supported by grants from the National Science Foundation (DUE-1224550) and the Alfred P. Sloan Foundation (2012627). The views expressed in this paper are solely those of the author and do not necessarily reflect those of the National Science Foundation and Alfred P. Sloan Foundation. In addition, Anne-Barrie Hunter, Mark Connolly, and Ross Benbow helped design the interview protocol and conducted many interviews and observations during the six site visits. Amy Mitchell Cowley assisted in developing the codebook and subsequent coding of the interview data. Other collaborators offered feedback on initial drafts of this article, including: You-Geon Lee, Julia Savoy, Elaine Seymour, Heather Thiry, Erika Vivyan, and Tim Weston. The author is also grateful to the three *LSE* reviewers and monitoring editor Marilyne Stains for their constructive criticisms and suggestions. All errors and omissions belong to the author.

## REFERENCES


Alexander, C., Chen, E., & Grumbach, K. (2009). How leaky is the health career pipeline? Minority student achievement in college gateway courses. *Academic Medicine*, *84*(6), 797–802.

Campbell, C. M., Cabrera, A. F., Michel, J. O., & Patel, S. (2017). From comprehensive to singular: A latent class analysis of college teaching practices. *Research in Higher Education*, *58*(6), 581–604.

Chang, M. J., Cerna, O., Han, J., & Saenz, V. (2008). The contradictory roles of institutional status in retaining underrepresented minorities in biomedical and behavioral science majors. *Review of Higher Education*, *31*(4), 433–464.

Coburn, C. E. (2001). Collective sensemaking about reading: How teachers mediate reading policy in their professional communities. *Educational Evaluation and Policy Analysis*, *23*(2), 145–170. https://doi.org/10.3102/01623737023002145

Corbin, J., & Strauss, A. L. (2008). *Basics of qualitative research* (3rd ed.). Thousand Oaks, CA: Sage.

Dancy, M. H., & Henderson, C. (2010). Pedagogical practices and instructional change of physics faculty. *American Journal of Physics*, *78*(10), 1056–1063.

Drinkwater, M. J., Matthews, K. E., Seiler, J., & Smith, M. (2017). How is science being taught? Measuring evidence-based teaching practices across undergraduate science departments. *CBE—Life Sciences Education*, *16*(1), ar18. https://doi.org/10.1187/cbe.15-12-0261

Eddy, S. L., Converse, 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*(2)

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis* (5th ed.). West Sussex, UK: Wiley.

Feldman, A. (2000). Decision making in the practical domain: A model of practical conceptual change. *Science Education*, *84*(5), 606–623.

Ferrare, J. J., & Hora, M. T. (2014). Cultural models of teaching and learning: Challenges and opportunities for undergraduate math and science education. *Journal of Higher Education*, *85*(6), 792–825.

Formann, A. K. (1984). *Die latent-class-analyse: Einführung in die theorie und anwendung*. Weinheim, Germany: Beltz.

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

Gainen, J. (1995). Barriers to success in quantitative gatekeeper courses. *New Directions for Teaching and Learning*, *1995*(61), 5–14. https://doi.org/10.1002/tl.37219956104

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education*, *53*(2), 229–261.

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), 731–767. https://doi.org/10.3102/00028312040003731

Gibbons, R. E., Villafañe, S. M., Stains, M., Murphy, K. L., & Raker, J. R. (2018). Beliefs about learning and enacted instructional practices: An investigation in postsecondary chemistry education. *Journal of Research in Science Teaching*, *55*(8), 1111–1133. https://doi.org/10.1002/tea.21444

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.

Gower, J. C. (1985). Measures of similarity, dissimilarity, and distance. In *Encyclopedia of statistical sciences* (Vol. 5, pp. 397–405). New York: Wiley.

Halpin, P. F., & Kieffer, M. J. (2015). Describing profiles of instructional practice: A new approach to analyzing classroom observation data. *Educational Researcher*, *44*(5), 263–277.

Harwood, W. S., Hansen, J., & Lotter, C. (2006). Measuring teacher beliefs about inquiry: The development of a blended qualitative/quantitative instrument. *Journal of Science Education and Technology*, *15*(1), 69–179.

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.

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.

Holland, D., & Quinn, N. (1987). *Cultural models in language and thought*. New York: Cambridge University Press.

Hora, M. T. (2014). Exploring faculty beliefs about student learning and their role in instructional decision-making. *Review of Higher Education*, *38*(1), 37–70. https://doi.org/10.1353/rhe.2014.0047

Hora, M. T. (2015). Toward a descriptive science of teaching: How the TDOP illuminates the multidimensional nature of active learning in postsecondary classrooms. *Science Education*, *99*(5), 783–818. https://doi.org/10.1002/sce.21175

Hora, M. T., & Ferrare, J. J. (2012). *A review of classroom observation techniques used in postsecondary settings (White Paper)*. Washington, DC: Association for the Advancement of Science.

Hora, M. T., & Ferrare, J. J. (2013). Instructional systems of practice: A multidimensional analysis of math and science undergraduate course planning and classroom teaching. *Journal of the Learning Sciences*, *22*(2), 212–257.

Hora, M. T., & Ferrare, J. J. (2014a). 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., & Ferrare, J. J. (2014b). *The Teaching Dimensions Observation Protocol (TDOP) 2.0*. Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research.

Hora, M. T., & Hunter, A.-B. (2014). Exploring the dynamics of organizational learning: Identifying the decision chains science and math faculty use to plan and teach undergraduate courses. *International Journal of STEM Education*, *1*(1), 1–21. https://doi.org/10.1186/s40594-014-0008-2

Kane, R., Sandretto, S., & Heath, C. (2002). Telling half the story: A critical review of research on the teaching beliefs and practices of university academics. *Review of Educational Research*, *72*(2), 177–228.

Lotter, C., Harwood, W. S., & Bonner, J. J. (2007). The influence of core teaching conceptions on teachers' use of inquiry teaching practices. *Journal of Research in Science Teaching*, *44*(9), 1318–1347.

Lund, T. J., Pilarz, M., Velasco, J. B., Chakraverty, D., Rosploch, K., Undersander, M., & Stains, M. (2015). The best of both worlds: Building on the COPUS and RTOP observation protocols to easily and reliably measure various levels of reformed instructional practice. *CBE—Life Sciences Education*, *14*(2), ar18.

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*(1), 13. https://doi.org/10.1186/s40594-015-0026-8

MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. In *5th Berkeley symposium on mathematical statistics and probability* (pp. 281–297). Berkeley: University of California Press.

MacQueen, K. M., Mclellan-Lemal, E., Bartholow, K., & Milstein, B. (2008). Team-based codebook development: Structure, process, and agreement. In Guest, G., & MacQueen, K. M. (Eds.), *Handbook for team-based qualitative research* (pp. 119–136). Lanham, MD: Altamira.

Malcom, S., & Feder, M. (Eds.). (2016). *Barriers and opportunities for 2-year and 4-year STEM degrees: Systemic change to support students' diverse pathways*. Washington, DC: National Academies Press.

Marbach-Ad, G., Ziemer, K. S., Orgler, M., & Thompson, K. V. (2014). Science teaching beliefs and reported approaches within a research university: Perspectives from faculty, graduate students, and undergraduates. *International Journal of Teaching and Learning in Higher Education*, *26*(2), 232–250.

President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: U.S. Government Office of Science and Technology.

Prosser, M., Trigwell, K., & Taylor, P. (1994). A phenomenographic study of academics' conceptions of science learning and teaching. *Learning and Instruction*, *4*(3), 217–231.

Saldana, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). London: Sage.

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), 245–253.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: 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), 618–627.

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.

Sokal, R. R., & Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, *38*(22), 1409–1438.

Spillane, J. P., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, *72*(3), 387–431. https://doi.org/10.3102/00346543072003387

Stains, M., Pilarz, M., & Chakraverty, D. (2015). Short and long-term impacts of the Cottrell Scholars collaborative new faculty workshop. *Journal of Chemical Education*, *92*(9), 1466–1476. https://doi.org/10.1021/acs.jchemed.5b00324

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

Stes, A., & Van Petegem, P. (2014). Profiling approaches to teaching in higher education: A cluster-analytic study. *Studies in Higher Education*, *39*(4), 644–658.

Sunal, D. W., Hodges, J., Sunal, C. S., Whitaker, K. W., Freeman, L. M., Edwards, L., … Odell, M. (2001). Teaching science in higher education: Faculty professional development and barriers to change. *School Science and Mathematics*, *101*(5), 246–257. https://doi.org/10.1111/j.1949-8594.2001.tb18027.x

Suresh, R. (2007). The relationship between barrier courses and persistence in engineering. *Journal of College Student Retention*, *8*(2), 215–239.

Swap, R. J., & Walter, J. A. (2015). An approach to engaging students in a large-enrollment, introductory STEM college course. *Journal of the Scholarship of Teaching and Learning*, *15*(5), 1–21.

Teasdale, R., Viskupic, K., Bartley, J. K., McConnell, D., Manduca, C., Bruckner, M., … Iverson, E. (2017). A multidimensional assessment of reformed teaching practice in geoscience classrooms. *Geosphere*, *13*(2), 608–627. https://doi.org/10.1130/GES01479.1

West, E. A., Paul, C. A., Webb, D., & Potter, W. H. (2013). Variation of instructor-student interactions in an introductory interactive physics course. *Physical Review Special Topics—Physics Education Research*, *9*(1), 010109. https://doi.org/10.1103/PhysRevSTPER.9.010109

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.

Woodbury, S., & Gess-Newsome, J. (2002). Overcoming the paradox of change without difference: A model of change in the arena of fundamental school reform. *Educational Policy*, *16*(5), 763–782.

Yin, R. (2008). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: Sage.