Exploring Data-Driven Decision-Making in the Field:

How Faculty Use Data and Other Forms of Information to Guide Instructional Decision-Making

WCER Working Paper No. 2014-3
August 2014

Matthew T. Hora

Assistant Scientist
Wisconsin Center for Education Research
University of Wisconsin–Madison
hora@wisc.edu

Jana Bouwma-Gearhart

Associate Professor
Oregon State University

Hyoung Joon Park

Research Assistant
Wisconsin Center for Education Research
University of Wisconsin–Madison

WCER

Wisconsin Center for Education Research
School of Education • University of Wisconsin–Madison • http://www.wcer.wisc.edu/

Exploring Data-Driven Decision-Making in the Field: How Faculty Use Data and Other Information to Guide Instructional Decision-Making

Matthew T. Hora, Jana Bouwma-Gearhart, and **Hyoung Joon Park**

One of the defining characteristics of current U.S. educational policy is a focus on using evidence, or data, to inform decisions about a range of topics that includes resource allocation, teacher hiring and firing, and curriculum and instruction. Such an approach is viewed as a corrective to the way that teachers and administrators have made decisions in the past—on the basis of less reliable information sources such as anecdote or intuition—and is further viewed by advocates as a core feature of successful educational reform and improvement (Mandinach, 2012). Besides being used to improve organizational efficiency*, data-driven decision-making* (hereafter DDDM) is also being used to support and improve practice at the classroom level (Halverson, Grigg, Pritchett, & Thomas, 2007).

Yet research on data use in K–12 schools and districts has unequivocally demonstrated that the existence of data alone does not magically lead to improved teaching and student learning (Spillane, 2012; Coburn & Turner, 2012). This is not due simply to educator resistance; the voluminous data produced in today's educational systems are often not adequately diagnostic or meaningful to support effective change (Gill, Borden, & Hallgren, 2014). Indeed, research on data use indicates that a critical aspect of DDDM is the complicated process of translating raw data into useable information and actionable knowledge (Hamilton et al., 2009; Mandinach, 2012).

This has led many researchers to advocate for practice-based research on how educators notice, interpret, and construct implications about data in real-world settings in order to improve our understanding of how teachers actually make decisions in the field (Coburn & Turner, 2012). However, given the panoply of organizational, sociocultural, and personal factors that shape teachers' decisions and subsequent actions, understanding how and why educators use data to make instructional decisions also requires attention to context (Lattuca & Stark, 2011). Of particular importance for effective DDDM systems are the organizational tools and systems that facilitate the regular collection, analysis, and dissemination of information, and the cultural norms that define what constitutes valid data and govern routinized data-related social interactions (Spillane, 2012). With insights into how educators actually utilize data and what contextual factors support and/or inhibit effective decision-making, educational leaders can target those leverage points that are best suited to achieving the ultimate goal of improving educational practice (Spillane, Halverson, & Diamond, 2001; Coburn & Turner, 2012).

Yet little is known about how postsecondary faculty[1] think about and use data when making decisions about their teaching. While a focus on using data in higher education is evident in performance-based funding efforts (Tandberg & Hillman, 2013), attempts to create institutional rating systems (Kelchen, 2014) and learning analytics (Wright, McKay, Miller, & Tritz, 2014), to date no empirical research has been conducted on faculty use of data to inform

---

[1] By *faculty* we mean all people who hold undergraduate teaching positions—whether full- or part-time, in a tenure track or not—in postsecondary institutions, with the exception of graduate teaching assistants.

instructional practice. Promising lines of inquiry include research on classroom formative assessment techniques such as clicker response systems (Bruff, 2009) and the ways evidence showing the efficacy of interactive teaching techniques can convince faculty to improve their own teaching (Weiman, Perkins, & Gilbert, 2010), but these studies are more prescriptive than descriptive, and thus fail to illuminate in objective terms precisely how faculty think about and use data in real-world settings.

In this paper we address this gap in the literature by investigating how 59 faculty at three large public research universities used data as part of their course planning activities. In exploring this topic, we draw upon naturalistic decision-making theory, which focuses on the subtle dynamics between context and cognition in shaping how people actually make decisions in complex, ill-structured situations rather than on how they should think and act (Klein, 2008; Shattuck & Miller, 2006). Using an inductive approach to analyzing interview data, we found that in practice, faculty draw on a variety of information types—not solely numeric data—that vary according to their form, subject, and source. [2] This information is most commonly part of informal, private (i.e., one-person) reflection at the end of the term. In cases where data collection and reflection are more formally integrated into institutional routines, they are either viewed as inadequately relevant or diagnostic (i.e., student evaluations) or are limited to unique situations (e.g., team-taught courses or disciplines with accreditation mandates).

We suggest that these routinized patterns of data use (or lack thereof), while representing deeply entrenched cultural practices that should not be dismissed out of hand, may also inhibit informed instructional decision-making due to a lack of structure, time for reflection, and availability of high-quality information. Towards this end, we suggest that a need exists for higher education leaders to design policies and professional development initiatives that facilitate a more formal collection of and reflection about data by faculty. In pursuing such technical solutions, however, policymakers and educational leaders must carefully negotiate the tension between rigor and relevance (Zeichner, 2007), and learn from the challenges experienced in the K–12 sector regarding data use (Mandinach, 2012).

Background

While much of the literature on educators' data use focuses on K–12 settings, a growing body of work explores the use of data in postsecondary institutions. Much of this literature focuses on efforts to support decision-making at the institutional or college levels using decision support systems (Goyal & Vohra, 2013) and data mining and predictive modeling of large datasets (Baepler & Murdoch, 2010; Romero, Ventura, & Garcia, 2008). Many of these efforts can be seen as part of a larger trend in using these analytic techniques, but looming in the background is the growing pressure on colleges and universities to embrace a "culture of evidence" (McClenney, McClenney, & Peterson, 2007; Morest, 2009).

These pressures are evident in the Obama administration's recent emphasis on developing metrics for measuring institutional quality (Obama, 2013; Kelchen, 2014), and a

---

[2] In order to avoid terminological confusion, we use the terms "data" to refer to numeric information and "information" to refer to non-numeric forms, including narrative and verbal, information.

general push toward accountability in U.S. higher education by policy makers, accrediting agencies, and the public (Petrides & Nodine, 2005). While many of these arguments are currently rhetorical, in some instances DDDM is being used in ways that are not unlike the focus on compliance (with corresponding punitive measures) embodied in the No Child Left Behind (NCLB) and Race to the Top efforts. Perhaps the clearest example in higher education is performance-based funding, in which states allocate funding to public colleges and universities based on analyses of data such as student completion rates (Tandberg & Hillman, 2013). Taken together, the signs do indicate that higher education is poised to enter a data-focused phase not unlike that in the K–12 sector at the beginning of the 1990s.

Challenges to Using DDDM in Education

Yet the adoption of DDDM has not been without its challenges in educational settings. This is due in part to administrators' and teachers' resistance to data-use policies, based on their view that these efforts are heavy-handed attempts to control the profession of education through overly simplistic measures (Fullan, 2010) and that standardized assessment data are inappropriate measures for evaluating educational quality (Anderson, 2006). But perhaps the biggest challenge to DDDM is the fact that no matter how sophisticated or methodologically robust results are, data alone does not automatically result in high-quality decisions (Ikemoto & Marsh, 2007). The reasons for this state of affairs are both technical in terms of systems design, and socio-cultural in terms of ensuring that the resulting data are adequately relevant and diagnostic for teachers in the field (Gill, Borden & Hallgren, 2014).

For example, Mandinach (2012) argues that three elements must be present in order to facilitate the actual use of data by teachers and administrators: a robust data infrastructure, in-house staff with appropriate analytic skills or "pedagogical data literacy," and common cultural norms about the value of data use as well as organizational routines and policies that reflect these norms (see also Hamilton et al., 2009). While the technical aspect of information systems should not be overlooked, researchers are increasingly highlighting instructor sense-making and the critical role of whether data will become translated into actionable knowledge to inform decision-making. To facilitate this translation process it is essential for the data to be relevant and timely for educators and to provide diagnostics for the problem at hand (e.g., identifying struggling students), all of which should be supported by some form of technical assistance for teachers to begin fostering an ongoing process of reflection about their own teaching practice (Halverson, Grigg, Prichett, & Thomas, 2007). Thus, for teachers concerned with identifying how to prioritize instructional time, identify struggling students, and assess the effectiveness of specific activities.

Of course, these processes of data use do not unfold in a vacuum, but are deeply influenced by the sociocultural and organizational context of a school, college, or university. Data use in K–12 schools tends to occur in group settings where teachers and/or administrators work in teams which, in many cases, develop shared notions of what data mean (Spillane, 2012; Coburn, Toure, & Yamashita, 2009). Additional organizational factors that influence how these teachers use data include routinized practices and policies (e.g., monthly grade-level meetings), available time and resources for data analysis, and leaders who can create institutional norms and procedures that either support or inhibit data use (Coburn & Turner, 2012).

A similar body of work that pays close attention to how data are received, interpreted, and incorporated (or not) into institutional decision-making in higher education finds similar impediments to data use, including a lack of staff competent to effectively analyze data (Blaiche & Wise, 2010), little recognition of data collection and analysis as a professional responsibility (Blaiche & Wise, 2010), concerns regarding data validity and distrust of how data are used (Jenkins & Kerrigan, 2008; Petrides & Nodine, 2003), and inadequate ability to translate data findings/analysis into action (Briggs et al., 2003; Johnston & Kristovich, 2006).

Finally, evidence indicates that teachers draw upon a wide range of information—not just numeric data—when preparing for their classes. These other types of information include informal student feedback, homework assignments, and conversations with colleagues (Gonzales, Moll & Amanti, 2005). Indeed, Hamilton et al. (2009) suggest that these alternative forms of information should be incorporated as part of DDDM systems to complement the analysis of numeric data, particularly as a form of interim assessment. This raises the prospect that a sole focus on numeric data, regardless of their relevance to actual teachers in the field, may be misguided. Indeed, Mandinach (2012, p. 81) wonders if the "pendulum has swung far to one side" in focusing on "hard" evidence in favor of information that is relevant and responsive to the needs of educators in the real world (Mandinach, 2012, p. 81).

The Need for Practice-based Research to Frame and Focus Data-related Interventions

This body of literature on the challenges of DDDM sends a clear message: "Data does not speak for itself" (Coburn & Turner, 2012, p. 177). Instead, implementing effective DDDM systems are immensely challenging, due in part to the fact that designers must deal equally with the technical problem of creating robust information systems as well as the problem of providing data via these systems that are adequately relevant and diagnostic to teachers in the field. This prospect highlights the importance of understanding the types of information that faculty in real-world settings consider to be relevant to their work as well as the practices and contextual factors that either support or inhibit effective decision-making.

Research on reform implementation indicates that the mismatch between new policies or innovations and the realities of practice in local settings where they are introduced is a common reason for unsuccessful reforms (Rogers, 1995; Fishman, 2005). Educational reforms in particular are most effective when they support educators' growth while also respecting their existing knowledge, skills, and realities (Darling-Hammond & McLaughlin, 1995), and this holds true for research university faculty as well (Bouwma-Gearhart, 2012). Once identified, the practices of a group of educators can be used as a foundation for designing new interventions that are more closely aligned with the realities and constraints of existing practice (Cobb, Zhao, & Dean, 2009). Thus, the first step in designing and implementing data interventions should ideally be a diagnosis or description of local practice.

Second, the identification of data use practices can be viewed as potential "leverage points" that leaders can support or alter in order to facilitate optimal interactions and processes for teachers within their organization (Mandinach, 2012). But which leverage points to target in postsecondary institutions? The evidence suggests that routinized practices and the contextual factors that support (or inhibit) them are particularly important components of both individual- and organization-level change. Research suggests that when data practices are organized in a

deliberate fashion, they can facilitate conversations around data and improve instructional decision-making in ways that incrementally result in considerable changes to instructional practice (Coburn & Turner, 2012).

In this study, we adopt a naturalistic decision-making approach to providing robust accounts of how faculty think about and use data and other types of information (Klein, 2008). Naturalistic decision-making theory is based on the notion that in addition to some element of conscious deliberation, human decision-making is limited by certain principles of cognition (e.g., perception, memory capacity). An important component of this perspective is a belief that decision makers construct a simplified mental model of the task that allows them to manage the overwhelming perceptual inputs of the environment into a representation that is cognitively efficient (Klein, 2008). To explore subtle features of such decision-making processes, researchers in this area use an interviewing technique called the Critical Decision Method (CDM), which is a structured retrospective think-aloud whereby individuals walk through a decision-making process (Feldon, 2010). While other methods for studying educational practice exist, such as participant observation (Spillane, 2012) and semi-structured interviews (Halverson & Clifford, 2006), we found that the CDM was uniquely suited to exploring the subtle dynamics among context, cognition, and practice in real-world settings.

Methods

We used a descriptive case study design for this study, which involves an intensive analysis of a bounded unit that results in a rich and detailed account of a phenomenon (Yin, 2008). In this case, the unit is that of individual faculty as they make instructional decisions within their unique organizational context. Our research questions were as follows: (1) What types of data and other information are used by faculty? (2) How do faculty use data and other information to support their teaching practice? (3) What are some defining characteristics regarding faculty use of data and other information? (4) What constraints and/or affordances influence faculty use of data and other information in their teaching practice? We also conducted an in-depth case study of a single example of data and information use in order to bring these processes of data use to life.

Sampling

We conducted this study at three large, public research universities in the United States and Canada in spring 2013. As the study was supported by the National Science Foundation, the disciplines included were biology, geology, mechanical engineering, and physics. Research universities were selected for this study in part because of the large number of undergraduates being trained in science and engineering at these institutions. It is important to note that at each of the study sites, some sort of teaching reform initiative and/or policy innovation was underway that explicitly sought to foster data use among faculty. At institutions A and B, this intervention included hiring post-doctoral students who assisted faculty in creating formative and summative data systems for their courses. Another data-related initiative that affected all sites was in the form of accreditation pressures from both regional and discipline-specific (i.e., mechanical engineering) agencies. Finally, at Institution C a new "general education" curricular reform that mandated new data collection and reporting procedures from pertinent undergraduate courses was underway. Each of these efforts likely influenced the data reported in this study.

A non-random purposive sampling procedure was used to identify study participants. Faculty (hereafter called *instructors*) were included in the study population if they were listed as course instructors in each institution's course listings for the 2013 spring semester. We contacted 165 instructors via email to request their participation in the study, and 59 ultimately agreed to participate (36% response rate). Participants represented the following disciplinary groups: biology (n=19), mechanical engineering (n=12), geosciences (n=15) and physics (n=11), as detailed in Table 1.

Table 1. **Characteristics of instructor sample**

	Total	Institution A	Institution B	Institution C
Total	59	20	18	21
Sex				
Female	19	9	4	6
Male	40	11	14	15
Discipline				
Biology	20	10	5	5
Mechanical Engineering	12	3	5	4
Geoscience	17	5	5	7
Physics	10	2	3	5
Position type				
Lecturer/Instructor	26	11	6	9
Assistant Professor	7	1	3	3
Associate Professor	15	6	3	6
Professor	11	2	6	3

Data Collection and Analysis

The data collected for this study include the results of in-depth interviews with respondents. A team of four researchers collected all data. All researchers underwent the same training procedures, which entailed in-depth discussions about research instruments and their use in the field. In addition, the lead author observed and took detailed notes on a meeting for a team-taught course at which three instructors and a course coordinator planned for an upcoming set of classes. Observing this meeting was unplanned and was based on a late invitation by one of the interviewees. As a result of this serendipitous opportunity and the rich and detailed data that emerged from the meeting, we selected this group for the illustrative in-depth analysis included in this paper.

For the interviews with instructors, we followed the CDM approach by asking a focused question about a recently performed task (i.e., planning for a specific class), and then asking follow-up questions to delve more deeply into the decision-making process of each respondent (see Feldon, 2010). The key question focusing on the use of data for course planning was: "Tell me exactly how, if at all, you used any data in planning your next class." This question was followed by probes that examined in greater detail the type of data used, specific planning steps, and other contextual factors that influenced course planning.

Analyses of these data included three stages. First, we analyzed responses to questions about instructional planning using an inductive approach to qualitative data analysis (Bernard, 2002; Miles & Huberman, 1994). To identify the types of data used across respondents, two analysts independently reviewed a sample of five transcripts, and each developed a preliminary code list using an open-coding procedure in which words or short phrases were assigned to specific utterances. For example, the following segment of text was analyzed:

> I like to ask challenging clicker questions, especially when I'm introducing a new topic, and so I showed a video which had a bunch of misconceptions in it, and they were able to identify those misconceptions. It was a two-minute long clip of what people think about viruses and bacteria, and I know that students have some of those misconceptions. I didn't have to go look that up because I already know it from previous investigations of data.

After developing a preliminary code list from the transcripts, both researchers met to compare results. In this case, both analysts assigned the code "personal memory of data" to this fragment, while agreeing that "clicker data" should not be assigned because the respondent did not state that he used them for the class under consideration.

After extensively reviewing all preliminary codes, we developed a revised code list and re-analyzed a new set of five transcripts. At this stage, we used the constant comparative method, in which each successive instance of a preliminary code (e.g., personal memory of data) was compared to previous instances in order to confirm or alter the code and its definition (i.e., the constant comparative method) (Glaser & Strauss, 1967). After another process of revising the codes, we compiled a final code list, whereupon the first author then reviewed the entire dataset using these codes. The first pass revealed nine different types of information (e.g., numeric information from assessments, numeric information from student evaluations, verbal information from colleagues) that respondents utilized for their teaching (Research Question [RQ] 1).

Then, to address the remaining research questions, each transcript was reviewed to identify references to instructors' use of these data types when actively planning for a class. This was necessary because these references frequently occurred throughout the entire interview, and not only in relation to the question about data use. As part of this process, the analysts made notes in the margins of transcripts regarding data use practices; these notes contained a combination of verbatim quotes, summaries of the practice being described, and contextual factors influencing data use. These notes were used to create "data use summaries" for each individual. It was at this stage of the analysis that we noticed that certain underlying characteristics could be discerned regarding data use practices, such as the number of participants involved, the degree of formality for data practices, and so on. In response, categories for certain characteristics of data practices were added to the summaries. To assess the reliability of this stage of the analysis, two analysts prepared summaries for a sample of 10 transcripts and met to discuss the results. These summaries were then used to answer research questions about the data and other information use practices (RQ2), defining characteristics of data and other information use (RQ3), and contextual factors (RQ4).

Finally, the interview transcripts and meeting notes taken by the lead author were analyzed using the thematic network analysis technique. This technique is a structured approach

for identifying relationships between concepts in a graphic and time-ordered fashion (Miles & Huberman, 1994). The purpose of this analysis was to identify the distinct steps that made up the planning strategies or actions for group being studied. To identify these related steps, the interview transcripts and meeting notes were analyzed to identify explicit statements regarding relationships among the types of data, data use practices, defining characteristics, and constraints or affordances used by the group. The results were then compiled in graphic form.

Limitations to the study include a self-selected sample, the limitation of the interview method in its reliance on respondents' conscious awareness of how they use data in practice, and the confounding variable of existing data-related initiatives at the study sites.

Results

This section reports the data analysis results for each of the four research questions.

1. What Types of Data and Other Information Are Used by Faculty?

First, it is worth noting that for several respondents, the question about the use of data for their own teaching required additional elaboration by the interviewer. This is likely because for this population—STEM instructors for whom data are quantitative measures used for research purposes—the notion of "data" used in their disciplines does not translate well to an educational context. For example, one biologist responded to the question by asking, "Can you give me an example of the types of data you're thinking about?" While responses to such a question are not desirable due to the potential for influencing subsequent interviewee answers, the researchers offered examples including test results and student evaluations (both quantitative measures).

Interestingly, fifteen instructors responded "no" to the question about data use, but six followed up their statement of "no" by describing a variety of information sources that they in fact did use to inform their teaching. In these cases, it was clear that the respondent's answer was negative in relation to the traditional notion of data (i.e., numeric information) but that the instructor used various types of information in practice. For example, a physicist noted that she administered mid-term student evaluations:

> I can get some pretty useful feedback on things like too much text on your slides or going too fast, that you can actually change that make a difference for the next six weeks. So it's not actual data, because I don't ask them to rank issues. I just ask them to provide written feedback.

In this case, students' comments clearly did not meet her notion of what "data" really are. However, she did use the written student comments to inform subsequent decisions about her teaching. Given the focus of this study on illuminating instructors' practice "in the wild" of real-world settings, we decided to focus on all types of information cited by in response to our question about data.

For the instructors who clearly identified some type of data used for their planning purposes, we identified three distinct aspects of the information: its form, its subject, and its source. While it is the category of "source" that is most similar to the common conceptions of data (e.g., student assessments, student evaluations), we found it important to specify the form in

which the information was conveyed (e.g., numeric, verbal) as well as the subject of the information (e.g., student learning, course structure). These additional categories underscore how data can take different physical or ideational forms and also refer to completely different topics (see Table 2).

Table 2. **Form, subject, and source of data and other information used by instructors**

| Category	Form	Subject	Source
Data	Numeric	Student Learning	Student Assessments
		Student Experience in Class	Student Evaluations
		Student Learning/Teaching	Education Research
Other Information	Verbal	Student Experience in Class	Conversations w/Students
		Teaching	Conversations w/Colleagues
Other Information	Narrative	Course Structure	Curricular Artifacts
		Student Experience in Class	Formal Qualitative Methods
Other Information	Personal	Student Misconceptions	Personal Memory
		Course Content	Personal Memory

These results highlight the fact that the types of information used by instructors to inform their teaching extend beyond that of quantitative data, and that each of these information types varies according to three distinct characteristics (i.e., form, subject, and source).

2. How Do Instructors Use Data and Other Information to Inform Their Teaching Practice?

Next, the analysis revealed 10 distinct uses of data and other information that varied according to characteristics that included the form of information used, the timing of data use, and the nature of the subsequent decision (see Figure 1).

Data (i.e., numeric information) about student learning are collected via assessments (e.g., clicker questions) and analyzed in class to guide on-the-spot decisions about teaching (15). Fifteen instructors reported collecting and analyzing numeric student assessment data during a class period, often through clicker questions, pop quizzes, or worksheets. Based on how well students responded to the assessment, instructors quickly evaluated whether they needed to spend more time on the topic or could move on. For example, a biologist reported that she posed clicker questions to students as a way to review material from the previous class, and often could determine "when they just are not getting it" based on their answers, which then led to further discussion. This use of data requires a rapid evaluation of information in real time, with the results of the analysis being immediately applied to instructional practice.

Figure 1. **Number of respondents reporting use of data and other information to inform teaching practice (n=59)**
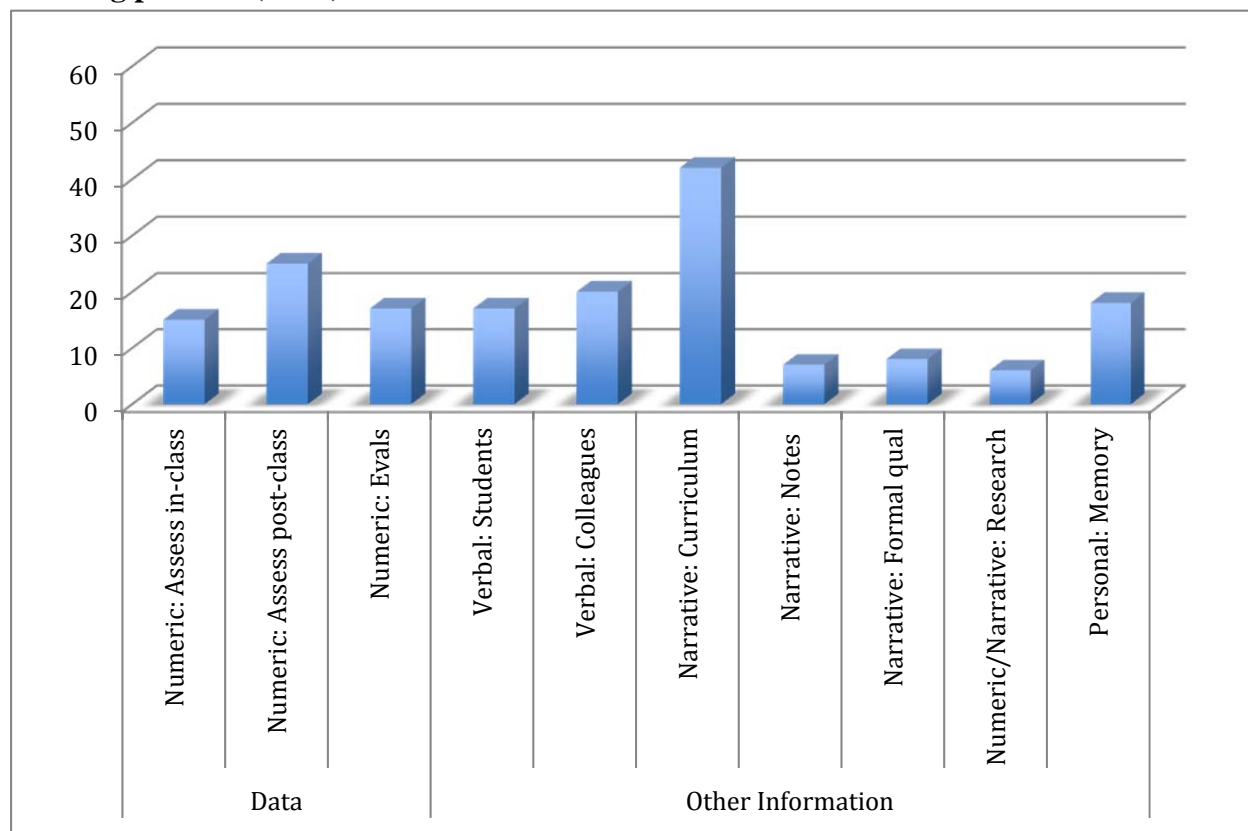
Data (i.e., numeric information) about student learning are collected via assessments (e.g., exams) and analyzed post-class or post-semester to guide decisions about teaching and/or course design (25). In addition to analyzing numeric student assessment data during class, 25 instructors also reported reflecting on collected data either after the class or at the end of the semester, ranging from mid-term exams to results from previous clicker questions. Analyses of these data then tended to guide decisions about how to teach a particular class and/or how the course itself should be structured the following year. A geologist made the following observation about this practice:

> I look at last year's clicker questions and say, "Oh, they did really poorly on this one. Let me try to improve." And I try to work on it right then but maybe if I remember from last year that they are not going to get this, I can actually teach it better before I ask the question.

In this case, the reflection on data from previous years led the instructor to identify topics that have proven to be challenging in the past, whereupon he could anticipate the next group of students having similar issues and adjusting his instruction accordingly.

Data (i.e., numeric information) about student experiences are collected via student evaluations and analyzed post-semester to guide later decisions about course design (17). The final manner in which numeric information was used by instructors was by analyzing student evaluations, which capture student perceptions about the quality of a course using both quantitative and qualitative measures. In this case, 17 instructors reported reflecting on the quantitative scores at the end of the semester to inform how that particular course would be designed the following year.

It is important to note that because institutionally supported student evaluations are only conducted at the conclusion of courses, these data could not be analyzed while the course was underway. One biologist noted that such data were "worthless" in terms of being applicable to her current course, which led her to create a new "satisfaction survey" that was administered in the middle of the semester. Since such surveys were discussed primarily in terms of qualitative data, instructors reporting them are not included in the 17 cases reported here, but are included with the group using formal qualitative data discussed below.

Verbal information about teaching and student experiences from students is recalled intermittently when making decisions about teaching and/or course design (17). In contrast to the type of data most commonly associated with DDDM (i.e., numeric student achievement data), instructors in our study sample reported that information about their own teaching as well as student experiences in their courses gleaned from conversations with them also represented a valuable source of teaching-related data. For example, a physicist noted that:

> I would say personal interactions weigh the most. Students that come to the office to ask for help, or students who just come for any reasons whatsoever, I usually quiz them: "What do you think is going on?" or "What did you learn in recitation today? What could change?"

Answers to questions such as these were used by 17 instructors in our study, both to diagnose specific problems in the classroom and to fine-tune their classes over the course of a semester.

Verbal information about teaching and student experiences from colleagues is recalled intermittently when making decisions about teaching and/or course design (20). Another form of verbally conveyed information used by a large number of instructors (20) was also about their own teaching and the experience of students in their courses, but here, the source was conversations with departmental colleagues instead of students. In some cases, this information was gleaned during formal post-semester meetings with other instructors in a team-taught course. In other cases, the sharing of information about effective teaching was somewhat informal. For example, one physicist noted that he and his colleagues "had a lot of discussions in the hallways, exchanging material from past semesters, (and generally) talking about teaching."

Narrative information about course structure and content is intermittently collected via curricular artifacts to guide decisions about course design (42). The most common type of information that instructors reported drawing on when planning and teaching their courses was narrative (i.e., written) information about their course contained in artifacts such as course syllabi and lecture notes. While this form of information is not typically thought of as "data" in

the same way as student achievement results are, this narrative information was central in informing instructors' decision-making. For example, one geologist noted that "I have the slides, lectures, and problems that I used in that class for the past three years—I use that as the basis for what I am going to talk about this year." In this way, these artifacts become perhaps the single most important source of information used to make decisions about how to teach a particular course.

Narrative information (about the success/failure of class activities) from personal notes on curricular artifacts is used to guide decisions about course design (7). In seven cases, the narrative information that instructors drew on while planning was their own notes taken the previous semester or year regarding what worked or did not work in a particular class. In these cases, notes were simply made on printouts of PowerPoint slides, scrap sheets of paper, or on lecture notes. For example, one geologist noted:

> I make notes (on the front of my notebook) at the beginning of the semester that say, "When I get to this topic, this is a change I want to make." And some of these notes are written now for next year already—things that I realize I need to do better.

In this way, notes made on existing curricular artifacts become part of an ongoing record of teaching tips that would be consulted in future semesters and years.

Narrative information about student experiences is collected via formal qualitative methods to inform decisions about teaching approach and/or course design (8). Several respondents discussed referring to narrative information about student experiences in the classroom while planning their courses. These data came in a variety of forms, including open-ended responses from student evaluation surveys, focus group reports or peer observation data (i.e., notes). In some cases, instructors created their own satisfaction surveys to administer at an early stage of the course as an improvement over the end-of-semester surveys used by most institutions. For example, one mechanical engineer reported that:

> I give an informal opportunity to the class members at week four to provide written feedback on how the class is working for them. If I do it about week four, I can get some pretty useful feedback. Things like, too much text on your slides or going too fast, going too slow. These are things that you can actually change that make a difference for the next six weeks.

Data (i.e., numeric information) and narrative information are recalled intermittently via educational research articles to guide decisions about teaching/course design (6). Another type of information used by six instructors came in the form of educational research articles, which contained both numeric (i.e., research data) and narrative (i.e., implications for teaching) information. In these cases, all six instructors drew upon published research papers, while three noted that they conducted their own educational research and actively drew on their own findings.

Personal information about course content and student misconceptions is recalled from personal memory to guide decisions about teaching/course design (18). The final type of information used by instructors was stored in their own personal memory. The memories and

knowledge most frequently described pertained to the content of the course and common student misconceptions. In one case, a respondent noted that she had taught the class five times and had developed an understanding of the "sticking points" for students, which led to an increase in the amount of time spent on these topics. In another case, the stored knowledge regarding a particular course was referred to as a "repertoire" from which the instructor drew.

3. What Are Some Defining Characteristics Regarding Instructor Use of Data and Other Information?

Once they had been identified, we sought to further examine instructors' practices to see if any patterns existed in their structure or overall composition. In particular, we were interested in certain aspects of data and other information use that the literature suggests are central to effective DDDM (e.g., degree of structure, number of participants in the system, nature of reflection, and role of students). With these characteristics of data use in mind, we re-analyzed instructors' practices and were able to further categorize them into more detailed sub-categories (see Figure 2 below).

Figure 2. **Characteristics of use of data and other information (n=59)**
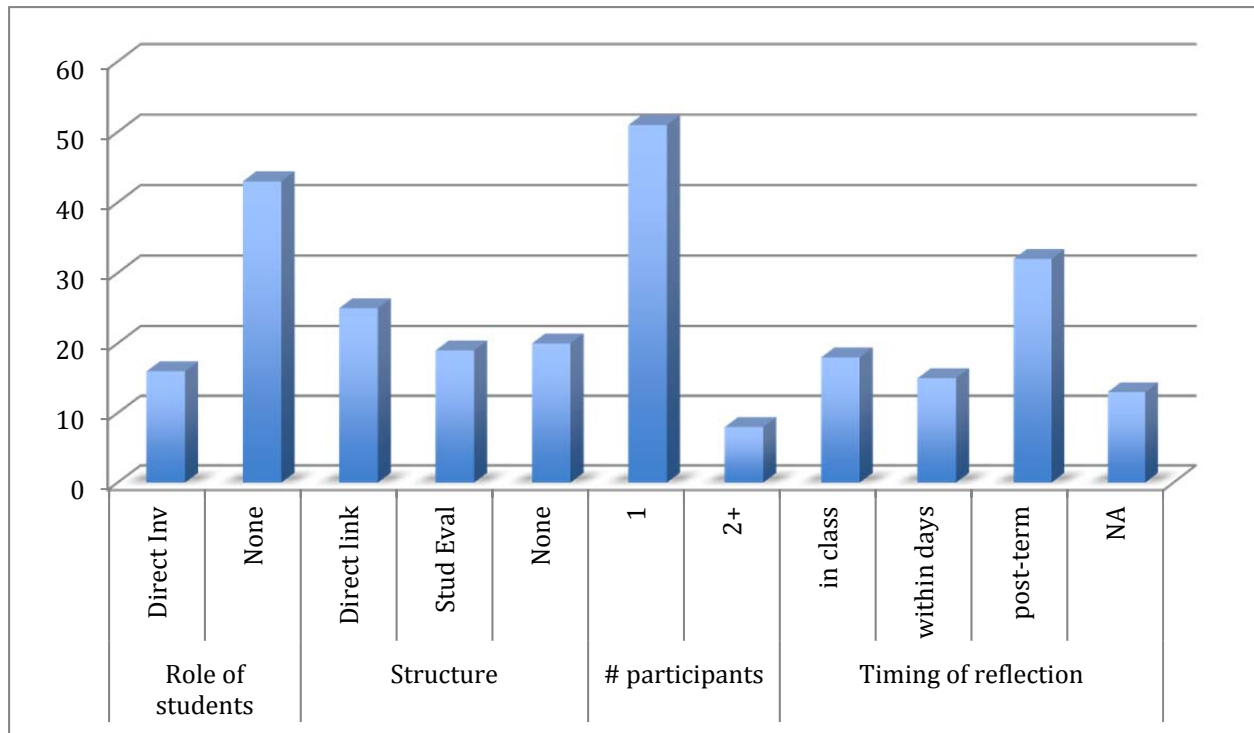
It is important to note that respondents were not directly asked about the relationship between their data and other information use and these characteristics; instead, they were volunteered in the course of the interviews. Thus, these results may under-represent actual practice.

Degree of structure. Research on DDDM systems converges on the need for a systematic, structured set of procedures and technology to be in place for the regular collection,

management, analysis, and interpretation of data (Halverson, Grigg, Prichett, & Thomas, 2007; Ikemoto & Marsh, 2007; Jenkins & Kerrigan, 2008).

No link to policy mentioned (20). Twenty, or roughly a third of instructors, did not mention any sort of policy or organizational structure related to their use of information for planning and teaching their courses.

Direct link to student evaluation data policy (19). Given the ubiquity of end-of-term student evaluations across all three study sites, we distinguish between policies governing the collection of these data from other data-related policies. For 19 respondents, the policy mentioned in relation to data use was that of their institution's end-of-term student evaluation. In these cases, the procedures in place for academic units to regularly collect these data using online or hard-copy questionnaires were mentioned as a part of organizational mechanisms that governed one type of data use.

Direct link to data-related policy (25). Twenty-five respondents discussed their data use in relation to a specific policy other than student evaluations. That is, some sort of mechanism was in place to support (or even mandate) their use of data for teaching purposes. Eight respondents mentioned accreditation policies as important policies governing data use, including institutional and discipline accreditation. In both cases, the fact that accreditation agencies are increasingly requiring data on student learning at the classroom or program level was noted. Seven respondents referred to program or departmental reviews, which took place on a 3 to 5-year cycle and involved institutional reviews of a program's (e.g., the undergraduate biology program) progress as measured by indicators such as graduation rates, student exit interviews, and so on. Finally, other mechanisms that shaped data use included institutionally mandated learning goals and courses with a fixed curriculum that required groups of faculty to coordinate their use of assessment and evaluation data.

Number of participants in data system. Much of the literature on DDDM assumes that the systems for collecting and analyzing data are multi-person: Teams of teachers and administrators are involved in the design and implementation of what are often rather complex technological systems. Nevertheless, there is no evidence that a "private" or one-person data system is undesirable or less than effective.

One participant in data system (51). Fifty-one respondents were the only person involved in collecting, analyzing, and interpreting teaching-related information.

Two or more participants in data system (8). Eight respondents used data in collaboration with two or more people. In each of these cases, the course under consideration was team taught, which required groups of instructors to work closely together to administer exams and manage subsequent data across sections throughout the term as part of a centralized system.

Role of students. Another potentially important component of instructional data systems is whether students are involved in the analysis of or reflection on data, which ultimately results in a form of formative feedback for students (Freed & Huba, 2000; Harrison et al., 2009).

Directly involved in reception/consideration of data (16). In cases where students were directly involved in the reflection on teaching-related data, they invariably occurred in the

classroom, where data from clicker-response systems were displayed on a screen for the class to scrutinize.

Uninvolved (43). In the majority of cases, students were uninvolved in the analysis and reflection on data. This is due to the fact that most instructors analyzed data in their own offices outside of the classroom.

Timing of reflection on data and other information. A core assumption underlying the effective use of data is that educators will pause to interpret the data and construct implications of the subsequent results for their own teaching practice (Coburn & Turner, 2012). Thus, some sort of reflection on the collected data is an essential part of translating data into actionable knowledge. In examining the data for evidence of some sort of reflective practice, we noticed that the timing of reflection was a key distinguishing factor among the instructors. That is, for those instructors whose descriptions of data use provided enough information on these points, they appeared to reflect on teaching-related data at three different points in time: in class, within days of class, or after the term was over.

In class (18). For 18 instructors, data were collected, interpreted, and quickly analyzed in real time, often leading immediately to an instructional decision in situ. For example, in cases where a large number of students answered incorrectly on a clicker question, one biologist reported that she typically changed her lesson plan mid-stream in order to spend more time on the topic. In these cases, the data "system" is compressed in time, though some preparatory work in developing meaningful questions and/or assignments is required.

Within days of class (15). Other respondents discussed reflecting on data within days of the class (e.g., Just-in-Time teaching, weekly quizzes, mid-term evaluations).

After the term is over (32). In many of these cases, instructors engaged in reflection at the conclusion of the course, examining the results of numeric assessments and student evaluations to make decisions about the next iteration of the course. For courses that were team taught, such as an introductory biology course at one of the study sites, this post-course process of reflection and evaluation was a departmental policy. One of the participants in these meetings described reviewing student assessment data, comments on evaluations and personal observations about successful and unsuccessful teaching activities from the semester. The group then made preliminary revisions to the course for the following year.

Not available (13). For 13 instructors, no details about the timing of their reflective practices (if any) were provided in the interviews.

4. What Constraints and/or Affordances Influence Instructor Use of Data and Other Information in Their Teaching Practice?

In this section, we report the various factors that instructors discussed in relation to their use of data and other information. It is important to note that whether factors acted as constraints or affordances to data use were most often analytic judgments made by the research team. In some cases, however, respondents did clearly state how a factor influenced data practices (see Table 3 below). Given the degree of inference used to compile these results, however, no frequency counts for each theme are provided.

Table 3. **Contextual factors that supported or inhibited the use of data and other information**

Supportive Factors	Inhibiting Factors
External accreditation policies	Lack of time due to workload
Policies for course, program, or departmental reviews	Poor quality of data
Local data-focused interventions	Timing of data
Institutional pressure to articulate learning goals	
Social networks supportive of instructional data use	

Supportive factors. A variety of factors appeared to support the use of data to improve teaching-related activities on the three participating campuses.

External accreditation policies. Instructors discussed accreditation criteria and procedures, either at the regional level where institutions as a whole are evaluated, or at the discipline level, where colleges, departments, and specific programs are evaluated, that act as external forces requiring administrators and faculty to collect, analyze, and report teaching-related data. This was particularly the case for engineering disciplines in the United States, where the Accreditation Board for Engineering and Technology (ABET) criteria include skills-based metrics that require faculty to collect and report data about student learning in specific competency areas (e.g., the ability to design and conduct experiments). In order to facilitate the collection of these data, ABET also requires departments and programs to articulate specific and measurable learning outcomes linked to each of these competencies. Further, ABET mandates that each academic unit demonstrate that it has instituted continuous improvement mechanisms, which are at the heart of many DDDM models, as follows:

> The program must regularly use appropriate, documented processes for assessing and evaluating the extent to which student outcomes are being attained. The results of these evaluations must be systematically utilized as input for the continuous improvement of the program. (ABET, 2014 Criteria for Accrediting Engineering Programs, p.4).

While the structures put in place by agencies such as ABET certainly do facilitate the regular collection and analysis of teaching-related data, what remains unclear is whether faculty at the department and/or classroom level actually reflect on these data and find these exercises useful or meaningful or treat them as simply a matter of compliance with accreditation requirements.

Instructors discussed similar mandates for collecting and reporting student learning data for their institutional accreditation. In this situation, the instructors felt that the actual learning outcomes for their discipline (i.e., physics) were determined by external agencies that had little knowledge of their field and their programmatic goals and requirements. Thus, the criteria were viewed as being imposed upon the institution in a top-down manner that was inimical to the

program. In contrast, while some engineers felt that the ABET requirements were onerous, they did feel that they accurately captured what their students should be learning.

Policies for course, program, and departmental reviews. In several departments at our study sites, instructors described formal procedures that governed collecting and reporting data in order to evaluate the quality of individual courses, degree programs, and entire departments. At the course level, these were often linked to team-taught courses for which instructors met weekly and at the course's end to review student assessment data and administrative functions of the course. These procedures were often informal in that they were implemented at the discretion of instructor(s). In other cases, formal program or curriculum reviews took place on a regular basis (e.g., every 3 to 5 years), and involved the collection and analysis of various forms of data (e.g., student achievement, student exit interviews) in order to assess the quality of a degree program. This sort of evaluation also existed for entire departments, although these procedures were often governed by institutional policy. In each case, policies for quality assurance essentially dictated the collection and analysis of data.

Local data-focused interventions. Another supportive factor for faculty data use is the existence of local data-focused programs or interventions. At two of the study sites, programs were underway that involved hiring post-doctoral students in STEM departments to assist faculty in articulating learning goals, developing formative and summative assessments to measure progress towards these goals, and perhaps most importantly, interpreting these data. As such, this program provided the human capital required to help translate raw data into actionable knowledge for teachers. Thus, these interventions were explicitly designed to support faculty DDDM.

Institutional pressure to articulate learning goals. Besides specific programs aimed at encouraging faculty data use, respondents at each study site described some form of institutional pressure (at the college or institutional levels) for faculty to begin articulating learning goals for students in their courses. In cases where these pressures were accompanied by penalties for non-compliance, as in the case of colleges of engineering that had to comply with ABET requirements on this point, the result was described as an "institutional culture" where data use was becoming the norm. In other cases, institutional encouragement to identify learning goals was merely seen as "pronouncements from the Provost's office" and carried little weight. In any case, the important role of leadership in creating institutional norms for data use is well documented (Coburn & Turner, 2012), and it was clear that DDDM is on the radar screen of institutional leaders at the three study sites.

Social networks supportive of instructional data use. Since skills and experience with collecting, managing, and interpreting pedagogical data is not common among research-oriented faculty, having colleagues and social networks that could be used to provide tips and suggestions on these points was viewed as an important factor. Such supportive networks for respondents tend to be closely related to those engaged in local data-focused interventions and other STEM education projects, whether at the department, institutional, or even national levels. In other cases, the social network referred to colleagues who were part of a team-taught course, where regular conversations and interactions surrounding teaching-related data were required.

Inhibiting factors. In addition, several aspects of the organizational context were reported to inhibit effective data use behaviors.

Lack of time due to workload. Respondents described their workdays as being filled with research, teaching, mentoring, and service responsibilities linked to promotion criteria, which in many cases led to 10–12 hour workdays. As a result, many felt there was little incentive to engage in a more rigorous approach to the use of, and reflection on, pedagogical data above and beyond what was mandated by their institution and department.

Poor quality of data. Another factor inhibiting the effective use of data is the perceived paucity of high-quality and useable data that instructors could use to inform their teaching. This complaint focused primarily on one type of data that many respondents felt could in fact help their teaching improve if it were of higher quality—end-of-semester student evaluations. Respondents noted that evaluations typically have low response rates and do not provide sufficiently detailed information about students' experiences to be useful. In this way, these data are viewed as inadequately diagnostic for specific aspects of student experiences or the course writ large.

The other commonly available form of data was that of instructor-initiated student assessments, and respondents generally did not raise questions regarding the quality of their own exams, clicker questions, or homework assignments. That said, several instructors noted that the ways in which these data were collected (i.e., as one-time summative scores solely for grading purposes) did not lend themselves to a continuous improvement process exemplified by DDDM. Thus, the limitation is less about the quality of the data itself, and instead concerns the ways in which they are integrated (or not) into an explicit procedure for continuous improvement.

Timing of data delivery. An issue related to the perceived poor quality of student evaluation data is the timing of its delivery back to faculty. The fact that reports are often sent back months after the conclusion of the course is viewed by respondents as a significant inhibitor to the effective use of the data. As one respondent noted, "By the time you get the student evaluations, you don't look at them for another year, or maybe you will never look at them again." At one institution, reports about the student evaluations were sent to faculty four months after the conclusion of a course, whereupon the next semester had already started, rendering the data "useless." Thus, these data are viewed by some as irrelevant due to their quality and lack of timely delivery.

5. In-Depth Analysis of Data Use: A Team-Taught Mechanical Engineering Course

While the results reported thus far illuminate key components of pedagogical data use, they are limited in their de-contextualization of practice from specific, concrete instances that is inevitable when data are aggregated across multiple cases. In this section we provide a brief yet in-depth analysis of a single instance of data use in order to bring to life how DDDM occurs "in the wild." In conducting the analysis we draw upon the results reported in the previous sections regarding types of data, data use practices and their key characteristics, and influential contextual factors to provide a conceptual framework with which to describe data use in action, as illustrated in Figure 3.

Figure 3. **Model of data use for ME 220 group**
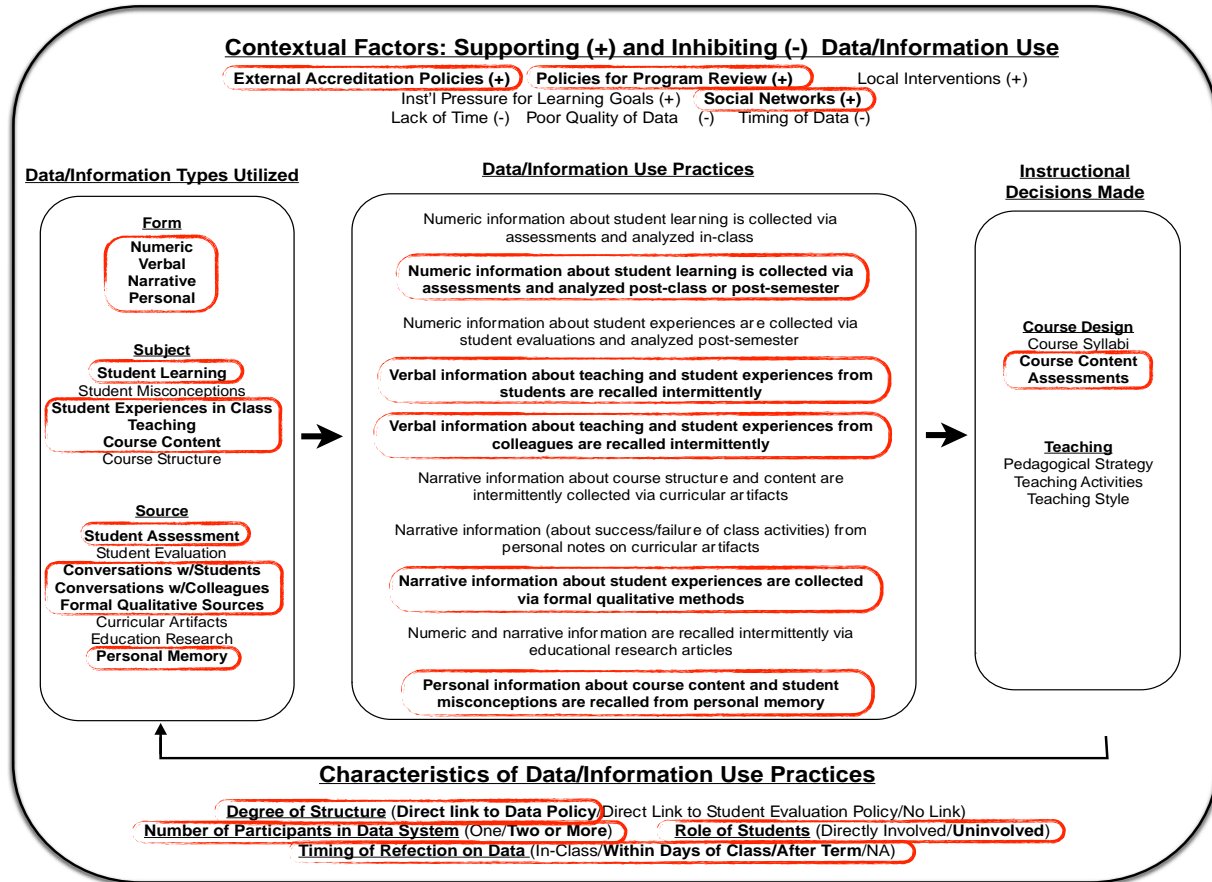
Our case is a team-taught course (ME 240) in mechanical engineering at one of the study institutions. This unique course was part of a set of three inter-connected courses in mathematics and engineering science that were specially designed for second-year students. A cohort of approximately 125 students each year goes through ME 240, attending classes four days a week as a single group. Three instructors (one mathematician and two mechanical engineers) and a course coordinator (a mechanical engineer) work in close collaboration with one another to ensure that the different class periods, assessments, and homework are all aligned with one another. Indeed, the classes that students sat through on any given day, which could be taught by all three instructors at different times, generally had some common themes linking the topics, homework, and assessments.

In order to effectively manage this extremely complicated course, the three instructors and course coordinator met weekly to discuss things like upcoming exams, results from the prior week's online quiz, and issues with particular students, as well as to reflect on conversations with students, numeric assessment data, and curricular artifacts from previous years: all to ensure that the coming week's classes and exams were coordinated and that each instructor was kept apprised of the cohort's progress. A variety of information in different forms, referring to

different subjects, and coming from various sources, was used to inform how this group planned and taught ME 240.

For example, narrative information from weekly student satisfaction surveys that contained complaints about an instructor's teaching style (i.e., too disorganized, too few example problems per class) was discussed. The group generally agreed that they liked these surveys but that response rate was a problem (lower than 10%) and that these particular complaints came too late in the semester to make adjustments. In any case, these observations were discussed in conjunction with an instructor's observations of tutorial sessions he had visited, as well as comments from teaching assistants about student satisfaction, as part of a larger conversation about how students felt about the course. This discussion did not clearly lead to a specific decision or action, as the information was not sufficiently timely or critical to merit an immediate change of course.

Next, the group talked about how to coordinate the discussion of Bernoulli's principle of fluid dynamics across each class period. One instructor reported how he had integrated the topic in his lecture, and another instructor noting that he had already identified a question for the final. This led to an extended conversation about the composition of the final, which began with an instructor pulling out a report of the statistical history of the exam in prior years and how many points should be allotted to each particular question. The coordinator observed that the questions included on Bernoulli's principle were not answered particularly well (i.e., average 37% correct answers). The author of those questions responded that he felt the students had not tried hard enough and that demonstrated mastery of the topic was essential for the second year of the program. The group then discussed other problematic questions including those on the ideal gas law, which appeared to have an error bar that extended below zero.

In summary, the reflection on numeric, narrative, verbal, and personal forms of information could be characterized in the following ways: highly structured (i.e., mandatory weekly meetings) with links to departmental policy involving two or more participants, not directly involving students, and taking place within days of a class as well as at the conclusion of the semester (via an end-of-year faculty retreat). Additionally, all these processes of data use, as well as the nature of the course itself, were described as being motivated by two issues: changes in the accreditation criteria that demanded more rigorous reporting of skills-based outcomes and internal dissatisfaction with the previous course sequence for second-year students. Taken together, these supportive contextual factors helped to craft a situation in which various types of pedagogical data were regularly reflected on in a highly structured and social environment. Finally, it is worth noting that this entire endeavor was based on what one participant called a "grass-roots" effort, meaning that faculty led the creation of the course sequence and its management (although the influence of accreditation mandates should probably not be underestimated). The group's members noted that they were aware of a data-focused intervention underway on their campus, but since it was not led by engineers, they felt no need to participate and also expressed skepticism that other scientists (e.g., physicists and geologists) would be able to assist them in a productive manner.

Discussion

In this section we discuss critical features of data use practices and assess the implications of these findings for policy, practice, and future research in this area.

What Constitutes Relevant and Diagnostic Data or Information for Faculty?

Developing effective DDDM systems requires an awareness of how educators in naturalistic settings actually use data and other forms of information to make decisions (Hamilton et al., 2009; Mandinach, 2012). To understand these processes of data use it is essential to identify those forms of data and other information that are considered salient, relevant, and diagnostic for faculty in their day-to-day instructional work (Coburn & Turner, 2012). As previously noted, much of the literature on DDDM and related educational policy assumes that the type of information being considered by faculty and "input" into organizational systems are numeric assessment data such as statistical analyses of student surveys (Pascarella & Blaich, 2013) or assessments of student learning (Weiman, Perkins & Gilbert, 2010). As the findings presented in this paper clearly demonstrate, however, the types of information utilized by faculty are far more diverse than such a narrow formulation suggests, and vary according to their form, subject, and source.

Anticipating the critique that these other forms of non-numeric information are not really "data," we reiterate that the focus of this paper is on the *information* that educators find credible and meaningful to their work, regardless of the form that it takes. One of the critical questions facing the field then, is whether some types of information are more valid, reliable, and meaningful than others, and in particular, who (if anyone) should be making such determinations? To answer this question, we argue that in terms of determining whether data are adequately salient and diagnostic to their own work, it is faculty themselves who should be the ultimate judge of a dataset's quality and utility. By prioritizing what practitioners experience and think over what external researchers consider to constitute high-quality data, or what cultural anthropologists call an "emic" perspective (Bernard, 2002), we thus ground our study in what faculty consider to be meaningful data on their own experiences.

For these instructors —STEM faculty in large research universities—one thing that is clear is that the notion of using data and other forms of evidence to inform instructional practice is not a distasteful or foreign idea. As one geologist noted, "We're in an evidence based field – we should be practicing what we preach." Along these lines, it appears that the "gold standard" of data for these faculty is numeric information about student learning, similar to the standard used in their own disciplinary research. This tacit understanding is apparent in the physicist quoted earlier, who viewed his mid-term evaluations as useful but "not actual data" because they included open-ended questions and not numeric (e.g., rank ordering) response options. As a result, it is clear that rigorously collected numeric data play a central role in not only common notions of DDDM writ large, but particularly for faculty in the STEM disciplines.

However, simply because data are in numeric form does not automatically make them useful. For example, one biologist who collected voluminous amounts of formative assessment data using clicker response systems felt overwhelmed with her data, and ultimately relied on her seven years of experience teaching the course and interacting with students to inform her teaching. While some proponents of DDDM in education would argue that this instructor should

not make instructional decisions based on such "intuition" and should instead collect more rigorous numeric data, we suggest that a myopic focus on such forms of information would overlook an important aspect of the craft of teaching. In practice another story emerges, in which a range of information types are considered as a basis on which to make decisions about course curricula, pedagogical strategies, and so on. It is clear that faculty—even those with highly sophisticated systems that utilize quantitative data—also use other sources of information that are meaningful in informing their practices. Thus, while the underlying motivation of DDDM in education is to replace non-numeric sources of information, such as anecdotes or intuition, it is important to recognize that, in practice, other sources of data play a critical role in how faculty think about their teaching, diagnose problems with student learning and their own instruction, and as the basis upon which to make future changes to their curriculum and instructional practice.

What Organizational Factors Facilitate Data-Driven Decision-Making?

Next, we focus on three factors that appear to support or otherwise afford faculty use of data and other information to support their instruction-related decisions.

Few structured opportunities for meaningful data collection and interpretation exist. First, we highlight the finding that for many of the faculty in our study, their use of data and other information to inform their course planning was limited to private, unstructured reflection that took place on an ad-hoc basis. Outside of team-taught courses such as ME 240 or departments where accreditation pressures mandated a highly structured approach to data collection, for many faculty the decision whether to collect, analyze, and utilize teaching-related data was left completely up to them. The main exception was that of student evaluations, whose collection is required in many institutions and supported by established systems for collecting and reporting data back to departments and instructors, but data that were largely ignored or deemed insufficient by our respondents.

While autonomy in one's teaching practice is of course a hallmark of academic life and work, the research suggests that effective DDDM requires not only robust technical systems but also social supports that facilitate the translation of raw data into useful knowledge (Mandinach, 2012). This is particularly the case for educators who are not skilled in analyzing and making sense of teaching-related data. As a result, some structured opportunities for the collection, analysis, and interpretation of data along with some professional guidance to translate these data into useable information, may be beneficial in postsecondary settings.

Little time exists (or is taken) for reflection about data. The notion of reflective practice is central to various theories of human development and learning, as well as organizational change and professional development (Schon, 1983; Kay & Johnson, 2002). As Larivee (2000, p. 293) argues, "Unless teachers develop the practice of critical reflection, they stay trapped in unexamined judgments, interpretations, assumptions, and expectations." Yet for many faculty in our study, this reflective process was unstructured and entailed brief glances at evaluation data or student final exam results. Again, notable exceptions include team-taught courses where groups of faculty held weekly and/or end-of-semester meetings to review student assessment data, evaluations, and to critically examine what went well (and what went wrong) with the course.

Reflection is ultimately how "raw" information, in its numerous forms, is translated into knowledge, and if this stage of reflective practice is missing, it is difficult to see how information of any type can be analyzed and its implications applied to improve real educational problems or tasks. Thus, for the instructors in our study who relied on their own memory as a source of information when planning their class, where this planning occurred 30 minutes before class on an ad hoc basis, it is difficult to argue that this type of "data or information use" could be considered an optimal form of DDDM. Similarly, the introduction of rigorously collected numeric data into academic departments, without corresponding time allotted for reflective practice, does not automatically equate with an evidence-based culture of teaching.

Of course, such reflective practice takes time, which is a resource that postsecondary faculty often find is in short supply. Time for reflection on teaching does not need to be burdensome, however, as the example of taking five minutes to make reflective notes after a class makes evident. While larger commitments of time such as daylong faculty retreats may be more effective, these may also require administrators to actively support, if not mandate, these activities. But without such a stage of critical reflection, even rigorous DDDM systems can too easily become another bureaucratic exercise that fails to improve educational practice.

The socio-cultural aspect of data and information use is critical. Finally, one of the facets of the organizational context that appears to strongly influence how data and other information are collected and used is that of the socio-cultural milieu in which faculty work. One of the striking aspects of the ME 240 case (as well as other team-taught courses in the study) was the social nature of the system in which regular reflection on data and other information types took place. Recall also that 20 respondents also discussed using verbal information, based on conversations with colleagues about student learning and teaching methods, when they planned and taught their classes. Thus, the value of opinions and experiences gleaned through social interactions with one's peers should not be underestimated.

It is important to recognize that while interactions such as these may occur naturally in the daily operations of an academic department, they can also be deliberately organized by leaders to facilitate certain routines and norms surrounding data use (Spillane, 2012). In instances such as regular meetings for team-taught courses like ME 240, colleague interactions serve as a structured opportunity to exchange opinions, experiences, and insights and facilitate reflection on this information. Otherwise, such exchanges are left to the happenstance of hallway encounters and water cooler conversations that may or may not lead to reflection and informed decision-making.

Implications for Policy and Practice

Using data to improve organizational efficiency and educational quality is becoming a key facet of educational policy in the early 21st century. Yet the evidence from the K–12 sector is clear. Based on over a decade of research since NCLB, it is clear that providing data alone is not the answer (Coburn & Turner, 2012). However, in the rush toward data mining, learning analytics, and institutional rating systems in the higher education landscape, Blaich and Wise argue that most leaders in colleges and universities assume that the problem of the effective use of data is technical and that "once we create sufficiently good measures, widespread institutional

improvement in student learning will follow" (2010, p.67). Instead, it is evident that the problem is not solely technical, but also social and cultural.

We argue that in designing policies and procedures that encourage DDDM among postsecondary educators, one of the critical points policymakers should consider is the fact that faculty utilize a variety of data and information in their daily work. Mandinach (2012, p. 81) raises the following questions on this issue:

> Education has often been accused of being a "soft" and unscientific field, thus the reliance on hard evidence and the emphasis on rigor. Has the field overreacted? Perhaps. And are educators being forced into overreliance on data? Perhaps. There needs to be a balance between the use of data and experience.

So, what do we propose? First, we draw attention to the importance of articulating a specific theory of change for academic organizations when thinking about implementing DDDM systems instead of simply introducing yet another innovation or policy change without a sense of the bigger picture. Educators have long questioned the efficacy of change efforts that solely involve mandating new practices from external authorities, or what can be considered a "top-down" approach. Such critiques of centralized reform are grounded in the recognition that such efforts often ignore the local realities of educational work and mandate too much change too quickly without adequate resources (e.g., Bouwma-Gearhart, 2012; Spillane, Reimer, & Reiser, 2002). However, an entirely grass-roots or "bottom-up" approach, where change is managed by local actors, may not be the solution either. Instead, a mix of both can be effective in harnessing the expertise, authority, and energy of both external authorities and practitioners on the ground (Fullan, 1994; Austin, 2011).

It is this integrative theory of change that we, as well as others studying data use in education (e.g., Coburn & Turner, 2012) suggest may be particularly well suited for educational reforms involving DDDM. In this study we discovered that external mandates governing data use, such as accreditation requirements and institutional procedures for program review, resulted in structured data collection and analysis activities that otherwise may not have occurred. At the same time, efforts to comply with these mandates that were designed "from the ground up" by faculty themselves, such as in the case of the ME 240 course, allowed instructors to craft unique responses to these external change efforts. Such an integrative approach is also desirable because policies and change initiatives mandated by external agencies (e.g., ABET) or institutional leadership will likely have a different impact on an institution and its faculty depending upon local factors such as the institutional and departmental cultures (Kezar & Eckel, 2002).

Based on evidence from this study, we suggest that two new policies that reflect this theory of change are worth exploring in the field: (1) Introduce mid-term student evaluations and provide feedback in a timely fashion, and (2) Mandate post-term faculty reflection on the course in the form of written reports that draw on these mid-term evaluations and other forms of data and information they find meaningfully informs their reflection.

These recommendations are based on the finding that at the present time, besides end-of-term student evaluation data (which are widely considered to be inadequately salient, timely, and

diagnostic), little information exists for faculty about how students are experiencing the course in general, and their teaching in particular. The provision of mid-term data on these points would give instructors real-time insights into these critical matters. Then, to spur reflective practice we suggest post-term reports – a recommendation that was inspired by a geologist in our study who made the following observation: "Our department chair does not require a written reflection from me at the end of the course – maybe he should." As we move forward with research in this area, we will be field-testing this exact approach in two large, research universities in the spring of 2015 by providing faculty with classroom observation and student evaluation data in the fifth week of the term (see Hora, Bouwma-Gearhart, Oleson & Collins, 2013 for an example of the data being provided to faculty; also see study website at http://tpdm.wceruw.org).

Besides attempting to maintain a middle ground between faculty autonomy and more structured departmental routines, these recommendations are based on the idea that specific leverage points exist within an organization that, when altered, may exert a significant effect on faculty work (Spillane et al, 2001). To be clear, we are not advocating an "everything goes" approach in which educator autonomy trumps the well documented need for teaching to become a more structured and data-based activity. Our research simply reminds us that we cannot ignore other sources of information that educators find useful and meaningful, and that we must also pay close attention to the social and structural components in which faculty practice unfolds on a daily basis.

References

Accreditation Board for Engineering and Technology (2013). Criteria for accrediting engineering programs. Retrieved from Accreditation Board for Engineering and Technology website: http://www.abet.org/uploadedFiles/Accreditation/Accreditation_Step_by_Step/Accreditation_Documents/Current/2014_-_2015/E001%2014-15%20EAC%20Criteria%203-13-14%282%29.pdf

Anderson, G. (2006). Assuring quality/resisting quality assurance: Academics' responses to quality in some Australian universities. *Quality in Higher Education*, *12*(2), 161–173.

Austin, A. E. (2011). *Promoting evidence-based change in undergraduate science education: A paper commissioned by the National Academies National Research Council Board on Science Education*. Retrieved from: http://www.tidemarkinstitute.org/sites/default/files/documents/Use%20of%20Evidence%20in%20Changinge%20Undergraduate%20Science%20Education%20%28Austin%29.pdf

Baepler, P., & Murdoch, C. J. (2010). Academic analytics and data mining in higher education. *International Journal for the Scholarship of Teaching and Learning*, *4*(2), 1–9.

Banta, T. W., Pike, G. R., & Hansen, M. J. (2009). The use of engagement data in accreditation, planning, and assessment. *New Directions for Institutional Research*, *2009*(141), 21–34.

Bernard, H. R. (2002). *Research methods in anthropology: Qualitative and quantitative approaches* (3rd ed.). Walnut Creek, CA: AltaMira Press.

Blaich, C. F., & Wise, K. S. (2010). Moving from assessment to institutional s improvement. *New Directions for Institutional Research*, *2010*(S2), 67–78.

Bouwma-Gearhart, J. (2012). *Engaging STEM faculty while attending to professional realities: An exploration of successful postsecondary STEM education reform at five SMTI institutions*. APLU/SMTI Paper 5. Washington, D.C.: Association of Public and Land-grant Universities. Retrieved from http://www.aplu.org/document.doc?id=4101

Bruff, D. (2009). *Teaching with classroom response systems: Creating active learning environments*. Hoboken, NJ: John Wiley & Sons.

Cobb, P., Zhao, Q., & Dean, C. (2009). Conducting design experiments to support teachers learning: A reflection from the field. *Journal of the Learning Sciences, 18*(2), 165–199.

Coburn, C. E., & Turner, E. O. (2011). Research on data use: A framework and analysis. *Measurement: Interdisciplinary Research & Perspective*, *9*(4), 173–206.

Coburn, C. E., Toure, J., & Yamashita, M. (2009). Evidence, interpretation, and persuasion: Instructional decision making in the district central office. *Teachers College Record, 111* (4), 1115–1161.

Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan*, *76*(8), 597–604.

Feldon, D. F. (2010). Do psychology researchers tell it like it is? A microgenetic analysis of research strategies and self-report accuracy along a continuum of expertise. *Instructional Science, 38*, 395–415.

Fishman, B. (2005). Adapting innovations to particular contexts of use: A collaborative framework. In C. Dede, J. Honan & L. Peters (Eds.), *Scaling up success: Lessons learned from technology-based educational innovation* (pp.48–66). New York, NY: Jossey-Bass.

Freed, J. E., & Huba, M. E. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning.* Needham Heights, MA: Allyn & Bacon.

Fullan, M. (1994). Coordinating top-down and bottom-up strategies for educational reform. In R. J. Anson, Ed.), *Systemic reform: Perspectives on personalizing education* (pp. 7–24). Washington, D.C.: U.S. Department of Education.

Fullan, M. (2010). *All systems go: The change imperative for whole system reform.* Thousand Oaks, CA: Corwin Press.

Gill, B., Borden, B. C., & Hallgren, K. (2014). *A conceptual framework for data-driven decision making* (No. 8142). Princeton, NJ: Mathematica Policy Research.

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

Gonzalez, N., Moll, C., & Amanti, C. (2005). *Funds of knowledge: Theorizing practices in households, communities, and classrooms*. Mahwah, NJ: Lawrence Erlbaum.

Goyal, M., & Vohra, R. (2012). Applications of data mining in higher education. *International Journal of Computer Science Issues*, *9*(2), 113–120.

Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist, 53*(1), 5–26.

Grimmett, P. P., & Mackinnon, A. M. (1992). Craft knowledge and the education of teachers. *Review of Research in Education*, *18*(1), 385–456.

Halverson, R. R., & Clifford, M. A. (2006). Evaluation in the wild: A distributed cognition perspective on teacher assessment. *Educational Administration Quarterly, 42*(4), 578–619.

Halverson, R., Grigg, J., Prichett, R., & Thomas, C. (2007). The new instructional leadership: Creating data-driven instructional systems in schools. *Journal of School Leadership, 17*, 159–193.

Hamilton, L., Halverson, R., Jackson, S. S., Mandinach, E., Supovitz, J. A., & Wayman, J. C. (2009). *Using student achievement data to support Instructional decision making. IES practice guide*. NCEE 2009-4067. Washington, D.C.: National Center for Education Evaluation and Regional Assistance.

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., & Wood, W. B. (2004). Scientific teaching. *Science*, *304*(5670), 521–522.

Hora, M. T., Bouwma-Gearhart, J., Oleson, A., & Collins, J. (2013). *Findings from an administration of the ISOP framework at Institution C in the Spring of 2013: Insights into course planning, classroom teaching, and student experiences in STEM courses.* Wisconsin Center for Education Research, UW-Madison.

Ikemoto, G. S., & Marsh, J. A. (2007). Cutting through the "data-driven" mantra: Different conceptions of data-driven decision making. *Yearbook of the National Society for the Study of Education*, *106*(1), 105–131.

Kay, J. K., & Johnson, K. L. (2002). Capturing complexity: A typology of reflective practice for teacher education. *Teaching and Teacher Education, 18,* 73–85.

Kelchen, R. (2014). *Moving forward with federal college ratings: Goals, metrics, and Recommendations*. WISCAPE Policy Brief. Madison, WI: University of Wisconsin-Madison, Wisconsin Center for the Advancement of Postsecondary Education (WISCAPE).

Klein, G. (2008). Naturalistic decision making. *Human Factors*, *50*(3), 456–460.

Larrivee, B. (2000) Transforming teaching practice: Becoming the critically reflective teacher. *Reflective Practice, 1*(3), 293–307.

Lattuca, L. R., & Stark, J. S. (2011). *Shaping the college curriculum: Academic plans in context (2nd Ed.)*. John Wiley & Sons.

Mandinach, E. B. (2012). A perfect time for data use: Using data-driven decision making to inform practice. *Educational Psychologist*, *47*(2), 71–85.

Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). *Making sense of data-driven decision making in education*. Retrieved from RAND website: http://www.rand.org/content/dam/rand/pubs/occasional_papers/2006/RAND_OP170.pdf

McClenney, K. M., McClenney, B. N., & Peterson, G. F. (2007). A culture of evidence: What is it? Do we have one?. *Planning for Higher Education*, *35*(3), 26.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks, CA: Sage.

Morest, V. S. (2009). Accountability, accreditation, and continuous improvement: Building a culture of evidence. *New Directions for Institutional Research*, *2009*(143), 17–27.

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107–110, § 115, Stat. 1425 (2002).

Pascarella, E. T., & Blaich, C. (2013). Lessons from the Wabash National Study of Liberal Arts Education. *Change: The Magazine of Higher Learning, 45*(2), 6–15.

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

Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, *51*(1), 368–384.

Schön, D. (1983). *The reflective practitioner: How professionals think in action.* New York, NY: Basic Books.

Shattuck, L. G., & Miller, N. L. (2006). Extending naturalistic decision making to complex organizations: A dynamic model of situated cognition. *Organization Studies, 27*(7), 989–1009.

Spillane, J. P. (2012). Data in practice: Conceptualizing the data-based decision-making phenomena. *American Journal of Education*, *118*(2), 113–141.

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.

Tandberg, D., & Hillman, N. (2013). *State performance funding for higher education: Silver bullet or red herring?* WISCAPE Policy Brief. Madison, WI: University of Wisconsin-

Madison, Wisconsin Center for the Advancement of Postsecondary Education (WISCAPE).

Usher, R., Bryant, I., & Johnston, R. (1997). *Adult education and the postmodern challenge: Learning beyond the limits*. London, England: Routledge.

Walsh, J. P., & Ungson, G. R. (1991). Organizational memory. *The Academy of Management Review*, *16*(1) 57–91.

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), 6–14.

Wright, M. C., McKay, T., Hershock, C., Miller, K., & Tritz, J. (2014, January-February). Better than expected: Using learning analytics to promote student success in gateway science. *Change: The Magazine of Higher Learning, 46*(1), 28–34.

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

Zeichner, K. (2007). Accumulating knowledge across self-studies in teacher education. *Journal of Teacher Education, 58*(1), 36–46.