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/

Hora. M. T, Bouwma-Gearhart, J., & Park, H. J. (2014). *Using Practice-based Research to Frame and Focus Pedagogical Reform: Exploring the Use of Data and Other Information to Guide Instructional Decision-making* (WCER Working Paper No. 2014-3). Retrieved from University of Wisconsin–Madison, Wisconsin Center for Education Research website: http://www.wcer.wisc.edu/publications/workingPapers/papers.php

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**

A defining characteristic of current U.S. educational policy at all levels is a focus on using evidence, or data, to inform decisions about institutional and educator quality, budgetary decisions, and what and how to teach students. This approach is often viewed as a corrective to the way that teachers and administrators made decisions in the past—on the basis of less reliable information sources such as anecdote or intuition—and is seen by advocates as a core feature of successful educational reform (Mandinach, 2012). Underlying the current push for data driven decision-making (hereafter DDDM) is the idea of continuous improvement, which refers to systems that are designed to continually monitor organizational processes in order to identify problems through careful analysis and then enact corrective measures (Bhuiyan & Baghel, 2005). In education this model has been widely adopted, particularly at the K–12 level, and is often associated with large datasets that are analyzed with sophisticated algorithms to identify which states, districts, and schools are succeeding or failing according to federal and state accountability criteria (Darling-Hammond, 2010).

Yet, research on data use in K–12 settings has demonstrated that the provision of data alone does not magically lead to improved teaching and learning. This is because DDDM is not solely a technical problem but one that involves translating raw data into information that administrators and teachers must find meaningful in their day-to-day work (Spillane, 2012). Imagine a principal and group of teachers struggling to understand precisely what voluminous amounts of student achievement data reports mean in terms of student advising, curriculum change, and classroom teaching. This process of integrating new information into pre-existing decision-making routines or creating new ones from scratch necessarily implicates the social networks, cultural norms, individual dispositions and information-processing tools and policies that collectively shape educators' behaviors within the complex organizations that are modern districts and schools (Halverson, Grigg, Prichett, & Thomas, 2007; Coburn & Turner, 2011).

Such insights have led to a robust line of inquiry on data use that is known as "practice-based research," which focuses on the how educators actually think, make decisions, and work in naturalistic settings rather than on describing the effects of interventions or proscribing best practices while ignoring real-world practices (Coburn & Turner, 2012). When applied to the use of teaching-related data, resulting accounts shed light on the organizational constraints and affordances that influence data use while also providing grounded accounts about how educators use data in their own organizations. These insights can then be used to inform the design of data-related interventions so that they are more aligned with the pre-existing norms and practices of specific organizations, rather than adopting a "top-down" approach that research demonstrates is a less than effective approach to reform (Spillane, Halverson & Diamond, 2001; Fullan, 2010).

So what does this all mean for higher education? Policymakers and postsecondary leaders devote considerable effort toward introducing a "culture of evidence" to higher education that is not dissimilar to the data-based accountability movement in K–12 education (Morest, 2009). This is evident in efforts such as performance-based funding (Hillman, Tandberg & Gross, 2014), institutional rating systems (Kelchen, 2014), and the increasing use of data mining and analytics to improve organizational operations (Picciano, 2012; Lane, 2014). At the classroom level, some argue that the use of large datasets and predictive modeling can improve teaching and learning through learning analytics, which is seen as an evidence-based way to identify struggling students, tailor instruction for different learners, and to generally improve faculty decision-making (Baepler & Murdoch, 2010; Wright, McKay, Miller, & Tritz, 2014). Taken together, these developments 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.

Thus, a pressing question facing higher education is whether the lessons learned from the DDDM movement in K–12 schools will be heeded. But, at the present time, little is known about the degree to which faculty utilize data and integrate them into their instructional decision-making processes. This is a considerable oversight given evidence that postsecondary institutions tend to focus more on the acts of generating and reporting data rather than applying the results to continually improve their practices (Blaich & Wise, 2010), a state of affairs that is exacerbated by the paucity of useful data about teaching that could be used for such purposes (Henderson, Turpen, Dancy, & Chapman, 2014). This suggests that the field lacks knowledge about the institutional and cultural environments into which innovations such as learning analytics and DDDM procedures will “land,” which is problematic given that there may not be a natural fit between them and the current way that colleges and departments are organized and the ways that faculty use data (or not) to make decisions about curriculum and instruction.

In this paper we address this gap in the literature by describing the data use practices of 59 science and engineering faculty from three large, public research universities as part of a larger research program aimed at bringing practice-based research to the study of faculty decision-making. In this qualitative, field-based study our goal was to document the lived experiences of these faculty in what cultural anthropologists call an “emic,” or insider, approach based on interviews using a technique from naturalistic decision-making (Crandall, Klein & Hoffman, 2006) and classroom observations. These data were analyzed using a combination of inductive thematic analysis, exploratory data reduction techniques, and causal network analysis in order to answer the following questions: (1) What types of data and other information are used by faculty? (2) What are some defining characteristics regarding faculty data use? (3) Can patterns be discerned in these data practices across the study sample? and (4) What role, if any, do these group patterns and local contextual factors play in shaping individual-level data use practices?

We discovered that faculty draw upon a variety of data and information in at least five distinct “repertoires of practice” or clusters of similar data-related actions that individuals (or groups) utilize in their work (Gutierrez & Rogoff, 2003). While some repertoires do not embody the principles of continuous improvement and many faculty continue to rely primarily on their intuition to inform their teaching practice, others have created sophisticated data systems to inform theirs in spite of the fact that current organizational structures are not amenable to DDDM. Ultimately, we argue that postsecondary leaders need to remedy this situation by providing better teaching-related data and creating incentives for and data-related policies about data use, but not at the expense of overlooking the diversity of existing faculty practices as well as the merits of experience gained through years of practicing their craft.

Background

DDDM has its roots in management, logistics, and business philosophies that view the regular analysis of and response to various forms of performance data as an essential component of organizational efficiency and productivity (Marsh, Pane & Hamilton, 2006). Techniques such as Total Quality Management and lean manufacturing are examples of systems whose core principle is that of continuous improvement, where feedback loops ensure that after a problem or inefficiency is identified, a new method can be tested and subsequent results integrated into new and improved procedures (Bhuiyan & Baghel, 2005). The role that data and other information play within an organization is central to continuous improvement, and related theories of organizations provide important insights into the relationships between data and change processes. In particular, theories such as information-processing (Galbraith, 1977) and organizational learning highlight the fact that data systems are comprised of both technical and sociocultural elements. This is especially true in relation to what is called an organization's "memory," or the mechanisms whereby data and other information are encoded, stored, and retrieved to inform decisions (Huber, 1991; Walsh & Ungson, 1994). Research on organizational change suggests that when alterations to these memory functions involve not just minor tweaks but instead transform central functions and cultural norms, it can be said that "organizational learning" has occurred (Argyris & Schön, 1978; Senge, 2006.

Data Driven Decision-making in K–12 Contexts

The ideas of organizational learning, continuous improvement, and DDDM spread far beyond matters of business and management to influence fields as diverse as medicine, government and public policy, and education. With the advent of high-stakes standardized tests in the 1970s, large datasets became available to make possible the implementation of continuous improvement in the field of education (Popham, 1987). In 2002, the US congress approved the No Child Left Behind (NCLB) legislation, which was seen by advocates as a remedy for decades of educational reforms in K–12 that had yielded little progress (2001). NCLB ushered in a new era of DDDM in education via federal policy that mandated statistical analyses of student achievement data to inform decisions about school funding, with ramifications for student retention, teacher hiring and firing and pay, school closures, and district organization (Darling-Hammond, 2010; Fullan, 2010). This legislation and its use of student achievement data as the key component of an accountability system serves as an important backdrop for the ensuing focus on DDDM across the educational spectrum.

DDDM in K–12 settings has been fairly well studied, and researchers have identified certain characteristics that comprise organizational data systems. These include the type(s) of data that are utilized to inform decision-making. While NCLB privileges outcome data, specifically student achievement as measured by standardized tests, other data has been utilized in K–12 data systems including input data (e.g. demographics), process data (e.g. teaching quality), and satisfaction data (e.g. of student, parents and other stakeholders) (Ikemoto & Marsh, 2007). Although some have made the case that other forms of information (e.g., homework assignments and student feedback) should be incorporated into data systems as a complement to numeric data, the term "data" almost always refers to numeric information in K–12 contexts (Hamilton et al., 2009). Other critical elements of DDDM include the technical infrastructure (e.g., computers, databases) for managing and reporting data, staff with "pedagogical data literacy" or expertise in analyzing and interpreting data for educational purposes, time and resources for educators to analyze and interpret data, leaders who create institutional norms or a "data culture" that supports data use as well as practices and policies that reflect these norms (Hamilton et al., 2009; Coburn & Turner, 2011; Liou, Grigg, & Halverson, 2014).

Of the myriad components that make up K–12 data systems, researchers have singled out two elements as particularly important features: the focus on continuous improvement, and the routinized practices of groups of educators. Unless the design of a data system integrates a cyclical process of data collection, interpretation, and application, the use of data can become a bureaucratic exercise in reporting numbers but with little substantive change in actual educational practice on the ground (Halverson et al., 2007; Mandinach, 2012). This cycle of feedback and continuous improvement is especially critical for leaders so that they can learn from prior experience and draw on data to intentionally redesign activities in order to achieve desired goals (Ikemoto & Marsh, 2007). Data-related practices also tend to involve a social component, as teams of teachers and/or administrators typically are engaged in collecting, analyzing, and constructing implications about the results for local matters. Importantly, it is within these regular meetings where groups develop shared commitments to data use and notions of what data mean (Coburn, Toure, & Yamashita, 2009; Spillane, 2012). As a result, Spillane (2012) argues that the routinized data-related practices of groups of educators, as a core feature of an organization's data culture and structure, are the grounds upon which data use interventions will succeed or fail (Spillane, 2012).

Researchers have also documented that administrators and teachers, with data at their disposal, engage in different types of DDDM that range from more to less effective, attributable to variations in the degree to which educators utilize data to inform practice (Ikemoto & Marsh, 2007; Marsh, Pane, & Hamilton, 2006). Ikemoto & Marsh (2007) argue that in practice, DDDM varies along two key dimensions, that of data complexity and sophistication of data analysis, both of which contain simple (e.g., single point-in-time data or descriptive statistics) and complex forms (e.g., trend data or inferential statistics). Based upon where particular data practice can be characterized based on these criteria, four main types of DDDM are proposed: basic, analysis-focused, data-focused, and inquiry-focused. Notably, the inquiry-based perspective of DDDM is characterized by the purposeful use of complex data analyses as part of cyclical feedback loops (Halverson et al., 2007).

DDDM in Colleges and Universities

While an extensive amount of research exists regarding how postsecondary institutions are organized (e.g., Bess & Dee, 2008), less attention has been paid to the specific role that data and data-related systems play in college and university operations. This is changing rapidly as the wave of accountability that engulfed K–12 schooling begins to influence postsecondary education, with growing pressure on colleges and universities to embrace a "culture of evidence" (Morest, 2009). These pressures are evident in the Obama administration's recent emphasis on developing metrics for measuring institutional quality (Kelchen, 2014), and a general push towards accountability in U.S. higher education by policy makers, accrediting agencies, and the public. 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 embodied in NCLB such as performance based funding, in which states allocate funding to public institutions based on data such as student completion rates (Hillman, Tandberg & Gross, 2014).

At the same time, given the extensive amounts of data available to most colleges and universities, such as graduation rates and annual tuition revenue, many observers argue that analytic techniques from the world of Big Data be applied to higher education (Picciano, 2012; Lane, 2014). In particular, at the classroom level there are now data produced by students via learning management systems and in-class clicker response systems that effectively represent "learner-produced data trails" (Long & Siemens, 2011, p. 32). As a result, it may be possible to use data mining and related analytic techniques that are now common in business and marketing to improve both institution-level operations (e.g., recruitment and admissions) and course-level decisions (e.g., tracking student progress) (Lane, 2014). The desire to incorporate aspects of DDDM in postsecondary classrooms is particularly evident in the science, technology, engineering, and mathematics (STEM) disciplines, where a considerable amount of resources are being invested into convincing faculty to adopt teaching methods that actively engage learners more so than traditional didactic lectures (PCAST, 2012). Some argue that one way to convince faculty is to produce evidence or data about the efficacy of these techniques (Weiman, Perkins, & Gilbert, 2010). In any case, an increasing amount of attention is being paid to the potential for data analytics to improve the nature of teaching and learning in the 21st century (Wright, McKay, Miller, & Tritz, 2014).

To date, however, little empirical research has been conducted on how faculty actually think about and use data as part of their instructional practice in real-world settings. An exception is a recent study that explored the basis upon which a group of biology faculty made teaching-related decisions, which found that personal experience and knowledge was utilized more than data or empirical evidence (Andrews & Lemons, 2015). The paucity of evidence regarding faculty data practices is problematic given that K–12 researchers have extensively documented the importance of organizational structures and sociocultural elements in shaping how educators use data. Yet, given the considerable variation between the nature of work in K–12 and higher education, including differences in incentive structures and the lack of common assessments to name but two, it is possible that DDDM would take different forms in a university as opposed to a K–12 school, thus indicating the need for empirical research on faculty data use practices.

Challenges with DDDM

The lack of practice-based insights is also problematic because research demonstrates that the adoption of DDDM has not been without its challenges in educational settings. At the core of the problem is the challenge of translating raw data into useable information and ultimately actionable knowledge that can be applied to real situations. Because of the necessity of this process, issues of local cultural norms related to data, individual motivation and habituated practices, and organizational routines come into play as an unavoidable aspect of data use. At the heart of the translation from data to knowledge is sense-making, or the process whereby individuals (and groups) notice certain types of data, interpret them in light of their own circumstances, and then draw implications for their own students or work (Coburn & Turner, 2011). Whether or not data are noticed and carefully considered depends on if they are perceived as adequately relevant, diagnostic, and valid by educators with respect to their immediate problems of practice (Halverson, et al., 2007; Gill, Borden, & Hallgren, 2014). But providing such data to educators is much easier said than done, and research in K–12 shows that the provision of data reports alone to principals and teachers is often viewed as the solution, with little attention to utility in the field (Spillane, 2012).

Researchers of data utilization in higher education have similarly found that getting educators to engage with data as part of their decision-making processes is rather challenging. In a study of factors shaping the adoption of data analytics in postsecondary institutions, Foss (2014) found that the adoption by academic leaders is shaped by a combination of features of the data system itself, the organizational context, and individual attributes of deans, chairs, and faculty. In particular, Foss (2014) found that, to be used, data must be viewed as legitimate within the profession and discipline, and useful for people in their daily work. Other researchers on data use in higher education have similarly found that cultural norms regarding data validity and utilization play a key role in dictating whether or not data will be considered, analyzed, and incorporated into regular practice (Jenkins & Kerrigan, 2008; Blaich & Wise, 2010). Finally, another challenge facing DDDM in postsecondary institutions is the inadequate capacity in regards to technology and human capital (i.e., skills) for translating data into useful and actionable knowledge (Johnston & Kristovich, 2000).

All these challenges with the use (or lack thereof) of data to enhance educational decision-making underscore the need for practice-based research, which is an approach that emerged, in part, as a response to overly normative research that examined why educators were not adopting educational "best practices" with fidelity, rather than exploring how people actually engage with new policies, curricula, or instructional techniques (Coburn & Turner, 2012; Spillane, 2012; Halverson & Clifford, 2006). A key component of practice-based research is a view of behavior as a situated phenomenon, such that activity is best understood not solely as the behavior of an individual but as their actions in relation to other actors (e.g., educators, administrators, and students), tools and technologies, and other contextual features within the "natural habitat" of real-world task situations (Hutchins, 1995; Greeno, 1998). Yet, the goal is not simply to describe activity at the micro-level but also to discern how organizational structures and cultural norms at the meso- and macro-levels interact with localized data practices.

What can be gained by a description of how faculty think about and use data? Such descriptive analyses can shed light on subtle aspects of behavior that provide valuable scientific insights in their own right, while also generating new hypotheses and theory for future research (Slavin, 2002; Flyvbjerg, 2006). Further, descriptive research would provide valuable insights into the current state of affairs regarding one of the most prominent educational reform policies of the early 21st century (Coburn & Turner, 2011). These insights can then be used to inform the design of programs and interventions in an approach similar to "user-based design," which is common in software development and industrial design and is based on the premise that adoption is dependent on the alignment between a new game, product, or tool and existing user behaviors (Spillane, Halverson, & Diamond, 2001; Fullan, 2010).

Methodology: Analyzing Characteristics of Data Use in Practice

The methodological approach utilized in this study is practice-based research, which can be located within the tradition of qualitative, field-based studies of naturalistic situations and settings (Coburn & Turner, 2012). In this paper we analyze semi-structured interviews and classroom observations using a variety of analytic techniques from within and outside the

DDDM literature, including thematic analysis, exploratory data reduction, and causal network analysis. The results shed light on the characteristics of data usage among all instructors in the study sample, and then how two of the instructors utilize data as part of their decision-making processes in specific situations.

Data Sources

This study was conducted at three large, public research universities in the United States and Canada, among disciplines that included science and engineering fields. Research universities were selected for this study in part because of the considerable efforts to transform undergraduate education at these institutions (PCAST, 2012). The three sites shared similar undergraduate enrollments and each had some sort of instructional reform effort underway that explicitly sought to foster data use among faculty. At Institutions A and B, this intervention (the Undergraduate Science Education initiative) included hiring post-doctoral students who assisted faculty in creating formative and summative data systems for their courses. At Institution C a general education curricular reform that mandated new data collection and reporting procedures was underway. Each effort likely influenced the data reported in this study.

A non-random purposive sampling procedure identified study participants in biology, geoscience, physics, and mechanical engineering. We selected these disciplines due to the large number of instructors across the study sites and for their leadership in educational reform initiatives. Faculty were included in the sampling frame if they were listed as course instructors in each institution's timetable. These courses included both upper and lower division courses such as Ecology and Evolutionary Biology, Introduction to Physics, Mechanical Component Design, and Environmental Geology. We contacted 165 instructors via email to request their participation, and 59 ultimately agreed to participate (36% response rate). Participants represented the following disciplinary groups: biology (n=19), mechanical engineering (n=12), geosciences (n=17) and physics (n=11). Faculty self-selected into the study and thus the results should not be generalized to the larger population of instructors at each site or in higher education (Table 1).

It is important to note that the percentage of instructors not on the tenure-track represented in the study (46%) was similar to the proportion of contingent faculty at participating institutions where such data were available (i.e., 33% and 47%). The course component of interest to the study was the in-class lecture period, and thus questions pertaining to laboratory and discussion sections were not included in the study protocols.

A team of four researchers conducted all data collection activities during the Spring semester of 2013. For the interviews we followed the Critical Decision Making (CDM) approach that improves upon traditional self-reports by using in-depth probes and think-aloud techniques to elicit respondent accounts about a specific recent activity (Crandall, Klein, & Hoffman, 2006). Prior to the beginning of each interview we reminded respondents that we were interested in a single course that they were currently teaching, and that all responses should be in reference to that course. The 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 contextual factors that influenced course planning. In asking these probes the interviewer encouraged respondents to be as precise as possible in articulating details about their use of data for the course under consideration. The remainder of the semi-structured protocol included open-ended questions about data-related topics such as continuous improvement efforts within departments. Thus, all respondents were asked the same questions, but interviewers were encouraged to pursue certain topics if they arose during the interview. Interviews took place in respondents' offices or nearby conference rooms and lasted approximately 45 minutes.

***Table 1*. Description of sample**

	Total	Institution A	Institution B	Institution C
Total	59	21	18	20
Sex				
Female	19	9	4	6
Male	40	11	14	15
Discipline				
Biology	19	9	5	5
Mechanical Engineering	12	4	5	3
Geoscience	17	5	5	7
Physics	11	3	3	5
Position type				
Lecturer/Instructor	27	11	6	10
Assistant Professor	7	1	3	3
Associate Professor	13	6	3	4
Professor	12	3	6	3

In addition, each of the participants was observed teaching one or two full class periods using the Teaching Dimensions Observation Protocol (Hora & Ferrare, 2013; Hora, 2015), which was utilized to code instructors' use of teaching methods (e.g., small group work), instructor-student interactions (e.g., types of Q&A), pedagogical strategies (e.g., humor), cognitive engagement (e.g., problem-solving), and instructional technology (e.g., clickers) at 2-minute intervals throughout a class period. Before collecting data, all four members of the research team underwent an intensive 28-hour training program over 2 weeks.

Identifying Characteristics of Data Use Practices

First, all interviews were transcribed and entered into NVivo qualitative data analysis software, whereupon two analysts (first and second author) segmented the raw data into more manageable and thematically coherent units (Gee, 1986). The segments pertained to three core topics that were central to the study: individuals' specific data routines for the observed classes, the existence (or not) of continuous improvement systems for their course, and contextual factors related to data use practices in general. Prior to segmenting the entire dataset, the two analysts first applied the topical codes to 10 transcripts and then compared coding decisions in order to ensure inter-rater reliability. Next, because it was evident that the text fragments remained rather complex and lengthy, an in-depth analysis of the data was conducted in order to prepare detailed summaries of each respondent's data use practices. These summaries distilled the raw data into short descriptions of the three core topics, using quotes and maintaining respondent language as much as possible. To create these summaries two different analysts (first and third author) prepared summaries of 10 respondents independently, met to compare results and made adjustments in order to arrive at a common understanding of what each summary should contain, whereupon the first author developed summaries for the entire sample.

We then developed a code list comprised of important features of DDDM as suggested by the literature, as well as an inductive analysis of the data summaries. The code categories selected for the analysis included types of data, types of data analysis, and types of continuous improvement. For each category we reviewed the literature and included themes that our respondents spoke to, such as the types of data suggested by Ikemoto and Marsh (2007) (e.g., input, process, and satisfaction). To complement these codes we conducted an inductive analysis of the data summaries using an open coding process to initially label interesting observations or ideas, and then each successive instance of the code 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). An example of a newly identified code for the types of data category includes real-time notes on teaching and direct feedback. This procedure of drawing on the literature and inductive analyses was repeated for the other two code categories.

The second stage then involved developing a participant by thematic code matrix in which each cell of the spreadsheet indicated whether participant *i* reported thematic code *j* (1) or not (0). To examine the degree to which the data practices captured in this matrix exhibited similarity or dissimilarity (i.e., underlying dimensionality), we used the exploratory data reduction technique of multi-dimensional scaling (MDS). The nonmetric MDS procedure utilized in this study graphically represents the similarity (or dissimilarity) between themes as distances in a two-dimensional space. For this analysis we used Euclidean distance to identify theme proximities. The procedure also provides a measure of the degree to which the resulting graph is consistent with a perfectly proportional graph of theme relationships, known as the "stress" value. Kruskal and Wish (1978) suggest that a cutoff for acceptable stress exists between 0.0 and 0.2 and the stress value for this analysis was 0.136. We also performed a hierarchical cluster analysis using Ward's Method in order to further explore the (dis)similarity of the themes and found a similar clustering of objects to the MDS analysis. Following these procedures, we returned to the data summaries in order to interpret the meaning behind the results. With the five groupings suggested by the analysis in mind, we were able to identify each "type" of data practice among the those reported at the individual level. Further, upon closer examination of the MDS results and the data summaries, we ultimately identified the nature of the differences among the five groups.

The third stage of the analysis involved an inductive thematic analysis of the text fragments for contextual factors influencing data use. This procedure included an open-coding process followed by the constant comparative method, whereupon a series of themes were identified that acted to either constrain or afford (i.e., support) effective data use.

The final stage entailed focusing on two instructors in order to examine the degree to which the five types of data use practices were evident at the individual level. The two cases were selected from the same study site in order to illustrate a range of data practices while holding the organizational context constant. In both cases the instructor's themes were closely examined to identify discrete chains of decision-making processes based on answers to the CDM question, as well as other salient factors that may have influenced these decisions (e.g., contextual factors). Additionally, each data practice reported was assigned to one or more of the five types of data practices previously identified. Then we analyzed the classroom observation data for each instructor by calculating the proportions that a particular code was observed in relation to all possible 2-minute intervals in the class period. This analysis also drew on a technique for combining codes to capture aspects of active learning (see Hora, 2015). Finally, we returned to the original transcripts to identify whether relationships could be identified among any of these data points. The determination of the existence of a relationship was based on respondents' explicit statements about associations between any two themes, factors, or behaviors. The results were then used to develop a graphic, called a causal network, that depicted the individual's entire decision-making process from planning to classroom instruction (Miles, Huberman, & Saldana, 2014).

Limitations to the study include a self-selected sample, the reliance of self-reported interview data on respondents' awareness of how they use data, and the confounding influence of existing data-related initiatives at the study sites. Also, given that the study sample reflects only a subset of disciplines from three research universities, we caution against generalizing the results to broader populations and encourage researchers to utilize the findings from this study to inform future studies with larger, more representative samples.

Results

1. Types of Data and Other Information Used by Faculty

First, it is worth noting that for several respondents the question about their use of data for teaching purposes 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—their notion of "data" does not translate well to their jobs as instructors. For example, one biologist initially expressed confusion at the notion that data would even be used in a teaching-related context. In another case, a physicist spoke broadly about data and information:

> 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, even though she did use written comments to inform subsequent decisions about her teaching, they clearly did not meet her notion of what "data" really are.

However, given our focus in this study on illuminating instructors' practice "in the wild" of real-world classrooms, we decided to focus on the broad range of potential data and information resources that faculty may utilize in their work. While this perspective departs from the notion of data primarily as numbers in the DDDM literature (as well as for our respondents), a broader perspective is consistent with our research goal of capturing the types of information considered salient and meaningful to organizational actors. As a result, in the remainder of the paper we discuss not only numeric data, but also qualitative data (e.g., open-ended survey responses), information gleaned from conversations with colleagues or the research literature, and direct feedback from students—as long as they were clearly identified as playing a role in how the respondent prepared for and monitored his or her teaching performance.

To categorize these different information types we adapted the framework proposed by Ikemoto and Marsh (2007) for types of data utilized in data systems (see Table 2).

***Table 2*. Types of data and other information utilized by faculty**

Data type	Example	# respondents
Input		
Personal experience	Memory about student misconceptions, course content, instruction-related issues	20 (34%)
Numeric data	Graduation rates, prior year's exam scores	11 (19%)
Colleagues advice/literature	Insights gleaned from conversations with colleagues and/or research literature	23 (39%)
Outcome		
Formative	Clicker question results, online reading quizzes	29 (49%)
Summative	Mid-term exams or finals	28 (47%)
Student satisfaction		
Institution provided	Standardized end-of-term student evaluations	29 (49%)
Instructor created	Unique mid-term evaluations or surveys	16 (27%)
Real-time notes on teaching	Instructor notes made about class	6 (10%)
Direct feedback	Student feedback delivered in office hours	16 (27%)
No data	No reference to any tangible form of data or information (could include personal experience and/or colleagues advice/literature)	12 (20%)

These data and other information types included input, outcome, student satisfaction, real-time notes on teaching, and direct feedback. The input category refers to those data and information that existed prior to the beginning of a course, whereupon the instructor retrieved them in the course of planning for the term or semester. The most common forms of input data were colleagues' advice, personal experience, and numeric data. The outcome category refers to those data that faculty gathered and utilized during the course regarding their students' performance on exams or other assessments, and include both formative and summative types. The student satisfaction category refers to data and information that captures students opinions about the course itself, and which took the form of institution-provided evaluations or feedback systems created by the instructor. Finally, two categories refer to types of data and information that respondents discussed collecting in the classroom itself regarding both student achievement and satisfaction—taking notes about how well a given class went (or not) immediately after class and also paying close attention to direct verbal or written feedback from students in the class.

2. What Are Characteristics Regarding Faculty Use of Data and Other Information?

Next, we analyzed our respondents' descriptions of how they actually used these data and information in practice. In doing so, we focused on certain aspects of data use that the literature suggests are necessary to provide a comprehensive account of data-related practices as well as topics identified in our inductive analysis of interview transcripts. These characteristics include types of analytic techniques, goals of data use, timing of data use and analysis, extent of participation with others, reliance on experts, frequency of analysis, application of data, and evidence of continuous improvement mechanisms (see Table 3).

In most cases respondents could report more than one type of data use (e.g., types of analysis). It is also important to note that for each category, we tabulated instances where either responses were not applicable or not provided. For example, in the case of the 12 faculty who reported no data use at all, tabulating their characteristics of data use was not possible. However, in other cases, respondents could have reported using data but failed to articulate in a precise manner certain characteristics of how they used them in practice. In the interest of space, these instances are not provided in Table 3.

Types of analysis. We identified three types of analysis—general, sophisticated, and reaction to feedback—based on the techniques instructors employed to analyze their collected data. The primary distinction between a “general” and “sophisticated” type of analysis pertains to the amount of time and degree of detail that an individual reported reviewing their data. For example, a mechanical engineering instructor who described simply “glancing” at exam and student evaluation scores, with no evidence provided regarding in-depth analysis of the data, was categorized as someone using “general” techniques. Such an approach was reported by 27 respondents (46% of the study sample). In contrast, analytic techniques considered to be “sophisticated” included a physicist who created scatterplots of students’ exam results at the end of the term and tracked results across multiple years, and a mechanical engineer who conducted a correlational analysis of the relationship between hours spent watching online tutorials and exam scores. This type of analysis was also reported by 27 respondents (46%). Interestingly, only eight respondents utilized both types of analytic approaches, thus indicating that faculty generally chose one or the other. Finally, the “reaction to feedback” category, which 14 respondents (24%) reported, included reports where faculty spent time reflecting upon student feedback (e.g., office hour conversations, in-class questions) and implications for their teaching.

Goal of data use. An important aspect of inquiry-driven DDDM (Halverson et al., 2007) is the articulation of goals for how data analyses will be used to improve instructional practice. To document this aspect, we identified two prevalent themes related to faculty goals for data use. First, 30 respondents (51%) discussed a goal of using data to “document student understanding” so that they could better understand and diagnose their students’ performance in the class. For instance, a biologist observed that data from clicker questions “showed me that they really didn’t get it as much as I had hoped (they) would,” which then told her that she needed to review the topic in the next class. Second, 43 respondents (73%) reported using data to “improve their course or curriculum.” In these cases, faculty utilized exam results to identify questions or problems that were too difficult or easy, which often led to the revision of assessments for the next semester. For instance, a physics instructor stated that he looked at the median value of exam scores which “helps me work out where it is too hard or too easy,” whereupon he adjusted the exam for the following semester.

***Table 3.* Characteristics of data use practices**

Characteristics of data use	Example	Total sample
Type of analysis		
General	Cursory review of exams or student evaluation scores	27 (46%)
Sophisticated	In-depth review of scatterplot, pre-post data, correlations	27 (46%)
Reaction to feedback	Reflection on student feedback	14 (24%)
No analysis reported		
Goal of data use/analysis		
Document student understanding	Goal to document student understanding/misconceptions	30 (51%)
Improve course/curriculum	Goal to apply results to improve course/curriculum	43 (73%)
No goals reported		
Timing of data use/analysis		
In class	Data are analyzed/applied during class in real-time	27 (46%)
Within days of class	Data are analyzed/applied after class but within days	23 (39%)
Post-semester	Data are analyzed/applied after semester is completed	44 (75%)
No timing reported		
Social nature of data use		
Solo	Data collection/analysis involves only 1 person	40 (68%)
Group	Data collection/analysis involves 2+ people	11 (19%)
No social aspect reported		
Reliance on data expertise		
Self-reliant	Respondent does not consult with data experts	37 (63%)
Consults with experts	Respondent consults with data experts	13 (22%)
Frequency of analysis		
One-time	Analysis occurs at single point during term or year	10 (17%)
Ongoing	Analysis of data ongoing throughout term or year	37 (63%)
No frequency reported		
Application of data		
Information applied	Application of results of data analyses to future work	44 (75%)
No application	No application of data analysis results mentioned	9 (15%)
Continuous improvement		
External policy (general)	Institution or disciplinary accreditation (e.g., ABET)	9 (15%)
External policy (USE project)	USE initiative focused on data use in STEM departments	14 (24%)
Internal policy	Departmental program review, student evaluations	42 (71%)
Personal CI systems	Respondent creates CI system on their own	31 (53%)
No CI systems reported		

Timing of reflection on data. A core assumption underlying instructional DDDM is that educators will pause to interpret the data and construct implications of the results for their own teaching practice (Halverson et al., 2007; Coburn & Turner, 2011). Thus, some sort of reflection on data is an essential part of translating them into actionable knowledge. In examining our data for evidence of reflective practice, we noticed that the timing of reflection was a key distinguishing factor among the instructors. Instructors appeared to reflect on teaching-related data at three different points in time: in class, within days of class, and post semester.

For 27 instructors, data were collected, interpreted, and quickly analyzed during class 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. For 23 instructors, data were reflected upon within days of a given class. In these cases data from sources such as weekly quizzes or mid-term evaluations were quickly analyzed shortly after results were available.

In 44 cases instructors engaged in reflection at the conclusion of the course, examining the results of 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.

Social nature of data use. Upon reviewing the data we were struck by a recurring theme that appeared to distinguish one type of data use from another—whether the respondent operated alone or with a group to collect and analyze the data. Interestingly much of the literature seems to assume that DDDM is occurring in group settings where teams of teachers and/or administrators work collaboratively. While there is no evidence in the literature indicating that a “private” or one-person data system is more or less effective, we documented the social nature of data use as an important characteristic of data use. In 40 cases, the respondent was the only person involved in collecting, analyzing, and interpreting teaching-related data and/or information. For 11 respondents, data were collected and analyzed in collaboration with two or more people. In each of these situations the course under consideration was team taught, and groups of instructors were required to work closely together to administer exams and manage subsequent data across sections throughout the semester as part of a centralized system.

Reliance on data expertise. The importance of staff who are skilled in analyzing and interpreting educational data is well documented in the literature. This is particularly important in organizations where instructors often lack pedagogical data literacy and/or the time to adequately analyze and reflect on results from analyses (Mandinach, 2012). As a result, we recorded the degree to which faculty in the study reported consulting with colleagues or staff who had expertise in educational data that they lacked. Of course, for the 37 who did not discuss such consultations, it is impossible to ascertain whether or not these individuals lacked such expertise. In any case, this large group of faculty relied exclusively on their own skills and knowledge to analyze and construct implications from collected data.

In contrast, 13 instructors consulted data experts to obtain assistance in activities that included assessment design, implementation of data-related technology (e.g., clicker systems), and the collection of original data to inform curriculum and instruction. In many of these cases, faculty reported consulting with staff hired through the USE initiative who served as local "data experts" within their department or college. For instance, one biologist reported that USE staff assisted him with conducting surveys of student satisfaction and knowledge, and to develop concept inventories for the course. Importantly, in this case these data were all input into an existing continuous improvement system (i.e., annual program review meetings) as a source of feedback focused on improving the first-year courses.

Frequency of analysis. Another theme that emerged from the inductive analysis of our data was the frequency with which respondents discussed analyzing their data. For 37 faculty data and other information were analyzed on an ongoing basis throughout the academic year. For instance, one physicist collected data using a concept inventory at the beginning and end of each course, and also analyzed student achievement data from quizzes on a regular basis. For 10 instructors, analysis took place only at a single point in time, often at the conclusion of the term or semester. While it is not possible to identify whether one-time routines of data analyses are more or less effective than ongoing ones, the latter do suggest that data are continually analyzed as an integral part of the individual's approach to curriculum and instruction.

Continuous improvement mechanisms. As previously discussed, at the heart of DDDM is a cycle of continuous improvement where mechanisms are put into place so that data can be regularly collected and analyzed to detect problematic procedures or activities, whereupon results are "fed back" into organizational operations so that corrections can be made. As a result, we analyzed the interview data to identify whether or not the presence of such systems could be identified. In doing so, we found that four types of continuous improvement mechanisms were in place. It is important to note that these results do not speak to whether or not corrections or improvements were made, but rather to the mere existence of some sort of feedback loop that influenced the respondent's program and/or course in some fashion.

First, 42 faculty reported internal organizational policies such as student evaluation policies or departmental program reviews, which often took place every 3–5 years, that influenced their courses to greater or lesser degrees. Next, 31 reported what we call "personalized continuous improvement systems," which means that these individuals had crafted their own course-level data systems involving the collection and analysis of data, followed by the application of results to their practice, without any institutional mandates or policy structures (see also Berger, 1997). Then, 14 respondents reported that the USE initiative had acted in some fashion to help with the design of continuous improvement systems. Finally, nine faculty reported that external policies (e.g., institutional accreditation) played a role in creating continuous improvement mechanisms.

Application of data to practice. Finally, one critical feature of DDDM is that the results of data analyses are actually applied in practice—the feedback loop component of continuous improvement systems. For 44 instructors the results of analyses were used to inform decisions about the course curricula and/or classroom teaching. Decisions related to the curriculum often involved altering future versions of the course (e.g., exams, content sequencing) based on the results of analyses, whereas instructional decisions generally impacted the style of teaching (e.g., pacing) as well as the time spent on particular topics. For instance, several instructors discussed using data from clicker questions to determine whether the next class should include more or less time spent on a particular topic. Of course, an interesting question is the degree to which continuous improvement mechanisms were related to the application of data to educational practice, an issue to which we now turn.

3. Do Patterns Exist in Faculty Data Use Practices?

The next phase of the analysis involved an examination of the patterns of data use across all faculty using scaling techniques. For this analysis we used MDS to explore the similarity among the types of data and information utilized by faculty, and characteristics of their data use practices. The result is a MDS graph that depicts the similarities among all variables in two-dimensional space, with those themes reported more frequently with one another across the sample clustered closely together (Figure 1).

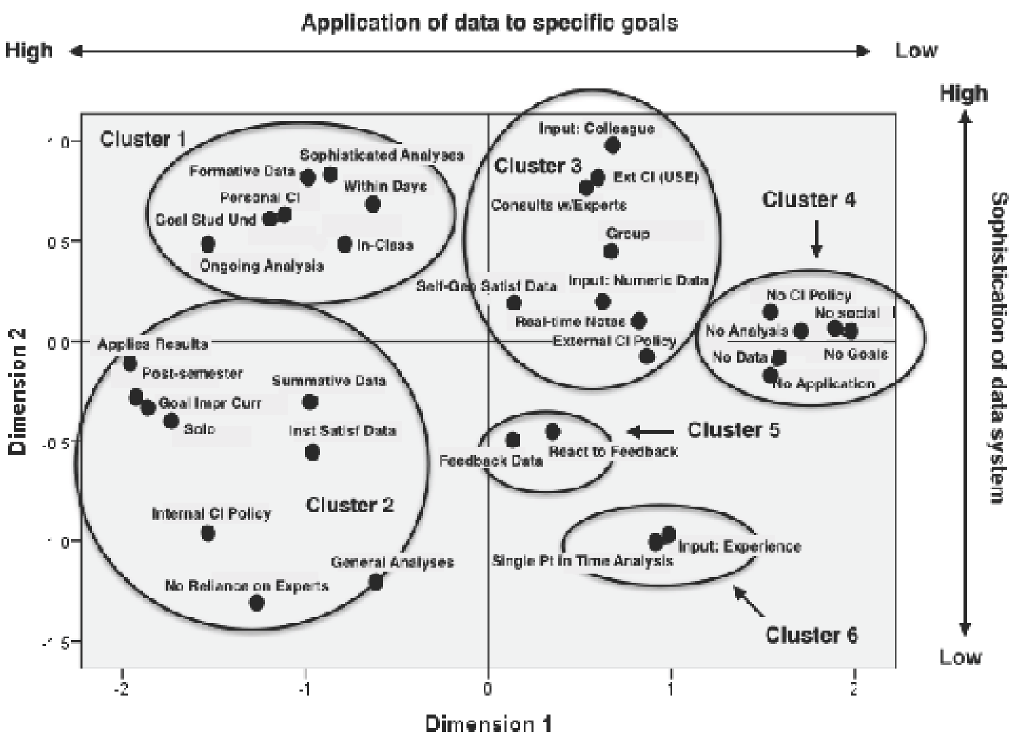

***Figure 1*. Multi-dimensional scaling analysis of interview themes**

Cluster 1 includes one type of data (i.e., formative outcome), one type of continuous improvement (i.e., personal), one goal (i.e., to document student understanding), one aspect of the frequency of analysis (i.e., ongoing), two aspects of the timing of analysis (i.e., in-class and within days), and one type of analysis (i.e., sophisticated). This cluster represents a set of practices initiated by the instructor herself that entails the sophisticated analysis of formative data either in-class or within days on an ongoing basis throughout the semester.

Cluster 2 contains a large number of themes including types of data (i.e., summative outcome and institution-provided student satisfaction), one aspect of the application of data (i.e., applies results), one type of analysis (i.e., general), one aspect of continuous improvement (i.e., internal policy), one goal (i.e., to improve curriculum), one aspect of the timing of analysis (i.e., post-semester), one aspect of the reliance on experts (i.e., none), and one aspect of the social nature of data use (i.e., solo). Taken together, these themes suggest a set of practices where individual faculty analyze, in a relatively cursory fashion, exam and student survey data after the semester is over with the goal of improving the next iteration of the course.

Cluster 3 also includes many themes, including types of data (i.e., input-colleagues, input-numeric data, real-time notes, and self-created student satisfaction), two types of continuous improvement (i.e., external: USE initiative; external: other), one aspect of reliance on experts (i.e., consults experts), and one aspect of the social nature of data use (i.e., group). These themes suggest a set of practices where faculty draw upon a variety of data in consultation with outside experts. These activities are either supported by external continuous improvement systems (e.g., USE initiative) or mandated by such systems (e.g., accreditation).

Cluster 4 includes each negative instance of data use, including no use of data, and the absence of data use characteristics, including goals for data use, analytic techniques, timing for data use, and so on. Taken together, these characteristics suggest a set of practices that are disengaged in any form of DDDM.

Cluster 5 includes one type of data (i.e., feedback) and one type of analysis (i.e., reaction to feedback), which suggest an approach to data use that relies on reacting to feedback obtained directly from students.

Finally, Cluster 6 includes one type of data (i.e., personal experience) and one aspect of the frequency of analysis (i.e., a single point in time). This cluster indicates that some faculty may reflect on their own experiences at a single point in time as a form of data utilization.

One advantage of MDS is that the grouping of variables is not the only story; the latent dimensionality of the distances must also be interpreted. In other words, the analyst needs to interpret the underlying rationale explaining why certain themes are arranged along the horizontal and vertical dimensions. We identified the nature of these dimensions through an iterative process of reviewing each respondent’s data summary in light of the results from the MDS analysis. The vertical dimension distinguishes among characteristics of data use based on the sophistication of data use as defined by advanced data analysis techniques and reliance on experts. The horizontal dimension distinguished among practices based on the degree to which the results of analyses were applied to specific goals or not, or what could be considered evidence of efforts to engage in continuous improvement.

The dimensionality of the graph suggests a more nuanced account of data practices than may at first be evident in the six cluster solution. First, the dimensions indicate that these clusters may represent types of data practices that can be viewed as being more or less inquiry-based depending on their degree of sophistication and evidence of continuous improvement efforts. While the results do not indicate a clear typology or ranking of data practices as suggested by other researchers (e.g., Ikemoto & Marsh, 2007), it appears that Clusters 1 and 3 are more sophisticated than the others, and Clusters 1 and 2 are more focused on application. While these results cannot be used to make claims about the subsequent quality of decision-making or instruction, it is safe to say that Clusters 4 and 6 represent practices that do not conform to the ideals of the DDDM movement. Second, the "spread" of many of the themes, particularly in Clusters 2 and 3, suggests that in practice these clusters may not represent clearly defined and mutually exclusive categories of behavior. Instead, they should be seen as collections of discrete practices that faculty may draw upon in their entirety, in pairs, or even singly. This suggests that a greater deal of variation may exist in the way that faculty enact various data-related practices, rather than individuals exclusively exhibiting the behaviors associated with a single cluster or "type."

4. Organizational Constraints and Affordances Influencing Data Use

Next, we report the various factors that instructors discussed as either constraining (-) or affording (+) their use of data and other information.

Lack of time due to workload (-). Respondents described their workdays as frequently exceeding 10 hours and being filled with research, teaching, mentoring, and service responsibilities. For faculty whose primary obligations was research, they often felt that there was little incentive for them to engage in a more rigorous approach to the use of and reflection upon pedagogical data above and beyond what was required by their institution (i.e., student evaluations). For faculty whose primary obligation was teaching, the workload was often sufficiently intense so as to limit the time available for engaging in DDDM.

Lack of expertise with educational data (-). A principal constraint facing effective data use is that most faculty, particularly in the science and engineering disciplines included in this study, lack expertise in working with educational data. The skills that respondents reported lacking included the ability to design and conduct educational research, analyze assessment data to identify patterns and construct implications, manage extensive amounts of varied educational data, and write effective assessments.

Poor quality of data (-). Another factor inhibiting the effective use of data is the perceived paucity of high-quality data provided by their institutions that instructors could use to inform their teaching. This complaint focused primarily on one type of data (end-of-semester student evaluations) that many respondents felt could, in fact, help their teaching improve if it were higher quality. Respondents noted that evaluations have low response rates and do not provide sufficiently detailed information about students' experiences to be useful. An issue related to the perceived poor quality of student evaluation data is the timing of its delivery back to faculty, which is often months after the conclusion of the course.

Course rotations (+/-). The common routine of rotating faculty into and out of teaching certain courses on a regular basis acts to both support and constrain effective data use practice. This is largely due to the fact that instructors typically design and accrue curricular artifacts (e.g., syllabi, exams, notes) over time that, when handed off to the next instructor in line, can represent a ready-made source of data. However, if the artifacts are neither well designed nor informative, the instructor has no prior data upon which to draw from and is then forced to start from scratch.

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). 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.

Policies for course, program, and departmental reviews (+). In several departments at our study sites, instructors described both formal and informal 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 on an informal basis to review student assessment data and administrative functions of the course. In other cases, formal program or curriculum reviews took place on a regular basis 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. In both cases, policies for quality assurance essentially dictated the collection and analysis of data.

Availability of local data experts (+). Another supportive factor for faculty data use is the existence of other faculty or staff who had expertise using educational data. These included networks of colleagues who regularly discussed discipline-based educational research, institution-based centers for teaching and learning, and funded projects focused on enhancing faculty data use such as the USE initiative. At two of the study sites, the USE initiative supported 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 interpreting these data. As such, this program provided the human capital required to help translate raw data into actionable knowledge for teachers.

5. Case Analyses: The Real-world Data Practices of Two Biology Instructors

The final step in the analysis was to examine how individual faculty enacted (if at all) the six different clusters of data practices as they went about planning and then teaching classes in the “natural habitat” of their institutions and departments. For this analysis we organized the specific ways in which two faculty utilized data to prepare for a specific class, the contextual factors that shaped these practices, and their classroom teaching practices into a causal network.

Dr. Robben. Dr. Robben was a full professor in the biology program at Institution B and was in his 11th year of teaching the lower-division course (Cellular and Molecular Biology). He reported three distinct data practices: (1) collecting student feedback data from a variety of sources (i.e., office hours, clicker questions, and in-class questions) that were then used to make changes in his exams and lectures, (2) making notes on PowerPoint slides that were used later to update the course, and (3) examining assessment results to identify problematic topics so he could emphasize them in later exams of lectures. Thus, the overriding concern for Dr. Robben in regard to using data to inform his decision-making was to identify student misconceptions and difficulties so that he could adjust his teaching accordingly. And while he drew upon a variety of data sources, he stated that "I would say personal interactions weigh the most."

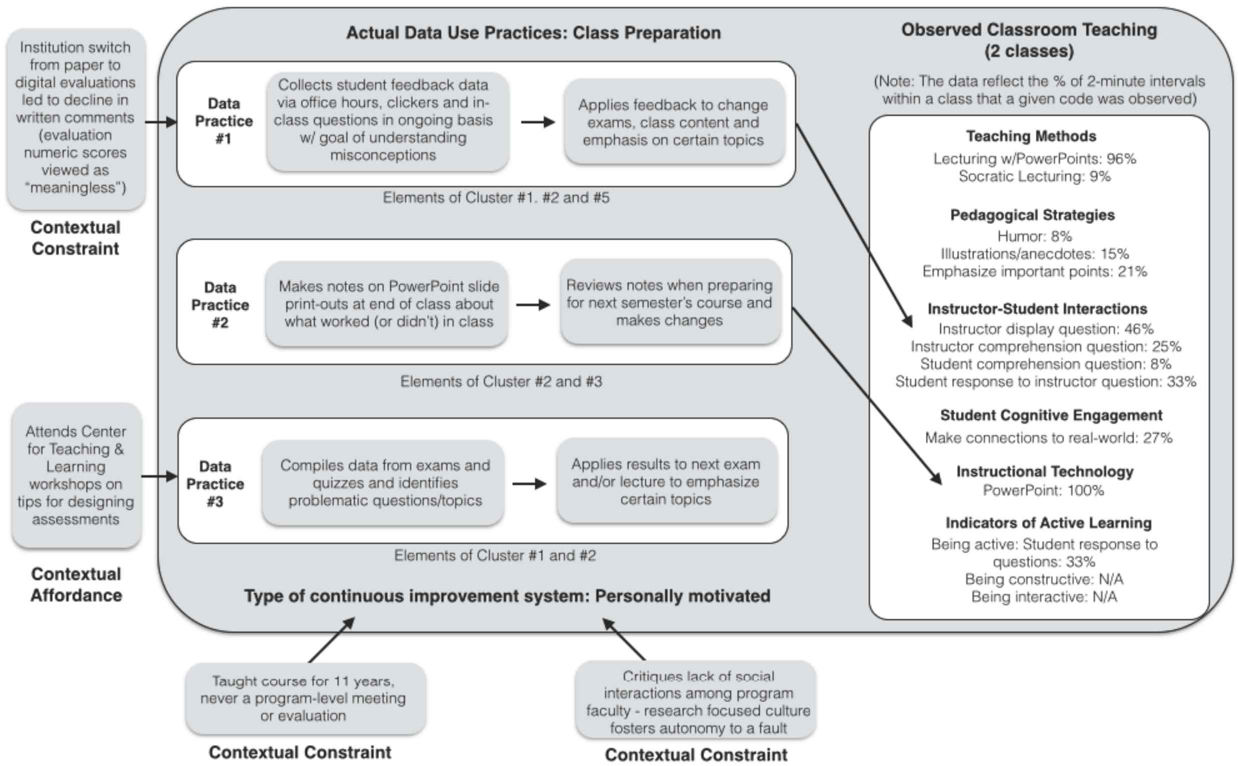

***Figure 2*. Causal network analysis of Dr. Robben's data practices**

These practices reflect aspects of data behaviors across Clusters 1, 2, 3, and 5, which indicates that while the clusters may represent regularities in data use across respondents, in practice an individual can draw upon multiple clusters in their daily work. Dr. Robben also reported that his data use was influenced by contextual factors including inadequate student evaluations, a useful center for teaching and learning, and the lack of social interactions and curricular reviews in the biology program. Ultimately, Dr. Robben's data use represented a "personal continuous improvement" system, and was evident in his classroom teaching through the use of regular questioning of students (observed in 46% of all 2-minute intervals) and the PowerPoint slides (100%) upon which he made notes after the class (see Figure 2).

Dr. Iniesta. Dr. Iniesta was a lecturer in the biology program at Institution A and was teaching a lower-division course titled Biology of the Cell. She reported two distinct data practices: (1) Using "just-in-time" teaching (pre-reading quizzes) to identify student misconceptions with the goal of then emphasizing difficult topics, and (2) reviewing end-of-semester student evaluations to identify problematic aspects of the course that were considered when preparing for the next semester. As with Dr. Robben, the primary concern with Dr. Iniesta was to identify student misconceptions and complaints so that her curriculum and instruction could be improved. However, Dr. Iniesta noted that while she had amassed an extensive database of online reading quiz data, she no longer referred to it because “the same misconceptions come up over and over again.” Furthermore, despite the assistance of data experts from the USE initiative helping her articulate new learning objectives and assessments and a desire to spend more time with data, she told us that “I would like to do more but I am overwhelmed.”

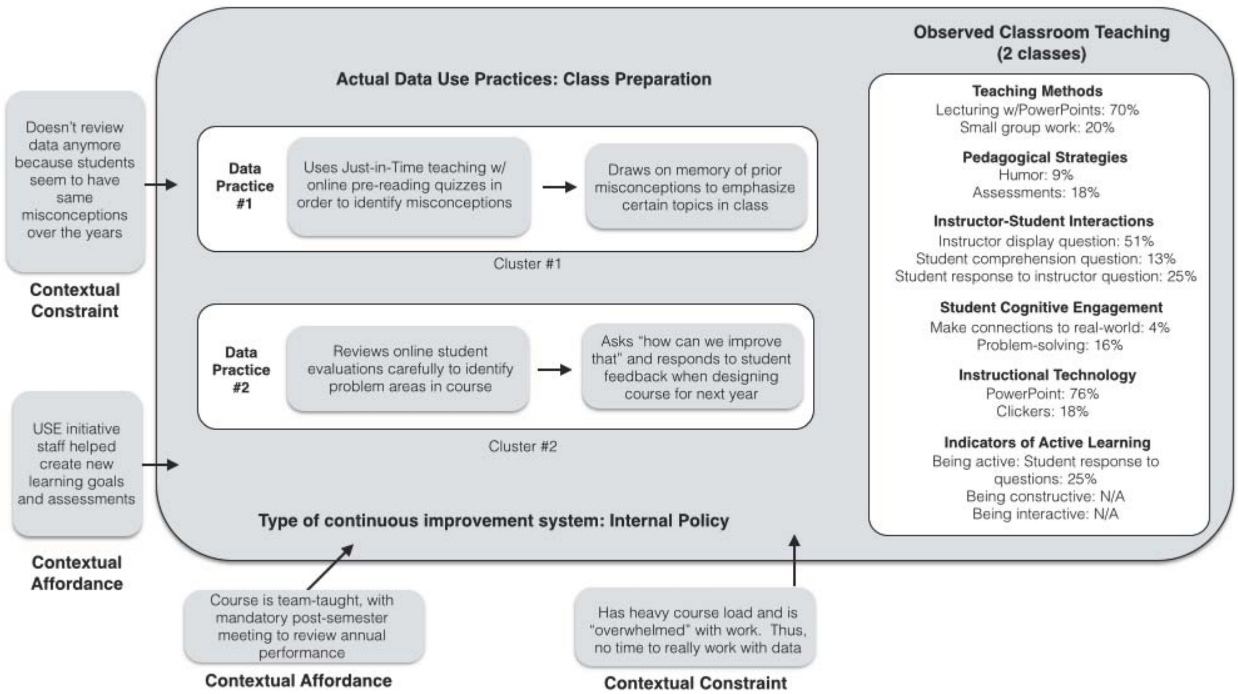

***Figure 3*. Causal network analysis of Dr. Iniesta’s data practices**

In any case, her reported data practices neatly fit within Clusters 1 and 2. Besides the intense workload, Dr. Iniesta also noted that because the course had multiple sections there was an annual post-semester review of the course where she met with the other instructors to review exam data, consider changes to the course, and so on. In this way, an internal policy for continuous improvement shaped some of her data practices. Finally, in the classroom Dr. Iniesta used a combination of lecturing with PowerPoint slides (observed in 70% of all 2-minute intervals), small group work (20%), and questions posed to students (51%) (see Figure 3).

Discussion

In this paper we reported the findings from a practice-based analysis of how a group of 59 science and engineering faculty think about and utilize teaching-related data in their “natural habitat” of real-world university classrooms. The results indicate that faculty drew upon six distinct repertoires of practice when using data and other information, with many engaged in the solitary collection and reflection upon data in organizations that are not designed to support DDDM. In this section we inspect key features of these practices and organizational contexts, and conclude with a critical analysis of the pros and cons of DDDM in higher education.

Are Faculty Engaged in DDDM? A Closer Look at "Data" and "Decision-making"

One pressing question facing higher education is whether or not faculty are presently engaged in DDDM as they prepare for and teach their courses. Our results indicate that for some faculty the answer is clearly no. This was evident in the 12 faculty in our study who referenced no data or other information whatsoever when speaking about their planning, and, more importantly, had no continuous improvement systems in place to inform their teaching outside of the ubiquitous end-of-semester student evaluations. Such an approach was captured by Clusters 5 and 6 in our analysis and was exemplified by the physicist whose idea of quality control was to informally track attendance in his courses as a proxy measure for quality.

Our data also indicate that there are many faculty who engage in some form of DDDM, but much depends on how one defines the term. Some faculty described formal, statistical analyses of numeric data as part of a continuous improvement process that, for some, represents the primary feature of effective DDDM system. Perhaps the best example of this was a team-taught mechanical engineering course where three instructors met weekly to analyze and discuss a variety of data (e.g., weekly quizzes, office hour conversations, in-class questions, mid-term results) to continually update their assessments and lecture topics. This conventional view of DDDM is implicit in efforts promoting the use of learning analytics and Big Data in higher education, as well as much of the accountability movement in the K–12 sector.

But there are many other instances where the use of data is less clearly aligned with these traditional expectations of DDDM. For example, Dr. Robben considered exam results as well as conversations from office hours and notes taken in real-time to inform his teaching as part of his own continuous improvement system. Given his reliance on not only qualitative data, but also ephemeral information (i.e., conversations with students) as the basis upon which to make his decisions, should his approach to decision-making be considered data driven? We argue that the answer is yes, but only because there exists evidence of a feedback loop wherein some form of evidence is carefully considered and then applied to the correction of a situation or problem. Consequently, besides those individuals whose practice lay solely within Clusters 5 and 6, we argue that the field adopt a broader perspective of DDDM that extends beyond the use of large numeric datasets and sophisticated algorithms to include other forms of data use practices.

This contention may be surprising to some readers, and to elaborate on our position we take a closer look at the two component parts of DDDM—that of "data" and also "decision-making"—two constructs that are often left unexamined in the literature. Upon closer scrutiny of precisely what these two terms mean, it becomes clearer that the distinctions between what constitutes DDDM and what does not are blurry at best, and more likely occupy a grey zone where such determinations must be made on a case-by-case basis.

Big and Small Data: The Dangers of Reifying Large-scale Quantitative Metrics

In recent years a backlash of sorts to the Big Data movement has become evident among practitioners who feel that the reliance on large datasets leads some to ignore "small data" sources such as interviews and surveys as well as the expertise of human beings whose knowledge is essential to contextualize and interpret the results of complex analyses (Peysakhovich & Stephens-Davidowitz, 2015). This position was famously made in the case of the "Google Flu" where analysts made incorrect predictions about a flu outbreak based on the results of a data mining exercise that drew on a dataset from Google Flu Trends. Unfortunately, the analysts failed to consider certain issues related to measurement and construct validity, but also did not incorporate more real-time health data comprised of smaller datasets available from the Centers on Disease Control in making claims about an impending pandemic (Lazer, Kennedy, King, & Vespignani, 2014). Cautionary tales about Big Data are not limited to the size of the dataset being analyzed or to particular statistical techniques, but also to the importance of the questions of “so what” or “why” in regard to a particular result. Consequently, as Peysakhovich and Stephens-Davidowitz (2015, p. 6) observe, Facebook’s data teams comprise “social psychologists, anthropologists and sociologists precisely to find what simple measures miss.” This point goes back to the finding in K–12 settings that data alone are not the answer to the complex issues facing education, but that human beings must then interpret the results, translate them to particular situations and contexts, and then determine appropriate corrective actions (Coburn & Turner, 2011; Mandinach, 2012).

Of course, this is not to minimize the benefits inherent in rigorous statistical analyses of large, high-quality datasets, which also conform to many postsecondary faculty’s views of what constitutes valid and reliable data (Wieman, Perkins, & Gilbert, 2010). Yet, the derivation and/or provision of these types of data alone are not sufficient, and other forms of information and types of analyses may also be necessary to answer complex questions. For instance, while some may argue that certain practices (e.g., reflecting on post-class notes made on PowerPoint slides) do not constitute a form of DDDM, we point out that for several faculty in our study these forms of information played an important role in how they continually improved their teaching practices. Other forms of “small data” include office hour visits with students, responses to open-ended questions on end-of-semester evaluations, focus group transcripts, and curricular materials that are “handed off” to faculty in a course rotation. Ultimately, policymakers should consider the fact that educators utilize a variety of data and information in their 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.

We strongly agree with this sentiment and encourage the field of higher education to embrace the role of big and small data, and quantitative and qualitative data in future conversations about how to use “evidence” and “data” to improve educational practice.

Decision-making: The pros and cons of expert intuition. Another aspect of DDDM that bears further scrutiny is the nature of decision-making that flows from (or not) the analysis of data. One striking finding was the strong reliance on personal experience as input data and the predominance of quick (i.e., in-class) solo, non-reliance on data experts for subsequent analyses. In other words, many faculty appear to rely on their expertise and intuition based on years of experience as a classroom teacher and/or as an expert in their disciplines. For example, consider the geologist who explained that because he had taught a certain class five times, “Yesterday I walked in through the door knowing what I was going to do and there was no formal planning process at all.” She had developed a strong understanding of students’ “sticking points” and thus felt no need to rely on data, or a slow, deliberate planning process. Also consider the case of Dr. Iniesta, who had amassed a database of reading quiz responses but no longer referred to them because she could anticipate student misconceptions beforehand. These findings that some faculty rely on “personal evidence,” which is often grounded in instructors’ perception that they have sufficient expertise to make sound decisions, have also been found by other researchers (Andrews & Lemons, 2015).

It is easy to dismiss decision-making like this as intuitive, anecdotal, and thus inferior to decisions made upon robust datasets and analyses. Certainly, it is undesirable for educators to engage in little to no reflection or consideration about their practice (Schön, 1983). But can we dismiss such decision-making out of hand? Research on expertise has shown that based upon thousands of hours of practice, a chess master can recognize between 10,000 and 100,000 separate patterns of piece configurations on the board, and recognize complex positions and layout of a particular game with a single glance (Simon & Gilmartin, 1973). Along similar lines, evidence suggests that experienced firefighters are able to make accurate decisions in the midst of crisis not through the use of deliberate, logical decision-making but through the rapid “search” through memory of particular cues and related responses (Klein, 2008). In the literature on DDDM there is an assumption that decisions made on the basis of intuition or an individual’s “hunch” are inferior to those informed by evidence, yet these examples raise questions about such a position. As the question by Mandinach (2012) above suggests, is it possible that an educators’ experience can lead to a robust knowledge of craft such that spreadsheets, pre-and posttests, and statistical analyses are not necessary to inform effective decision-making?

Also consider the issue of time, which is implicit in the above examples of expert decision-making, and the 27 faculty in our study who analyzed data on the spot in the middle of class, often in relation to student responses to clicker questions. This type of analysis or thinking is rather common in educational practice (see Shavelson & Stern, 1981) given the need for educators to react spontaneously to new and often unpredictable situations in the classroom. Such rapid decision-making can be thought of in terms of “reflection-in-action” as opposed to the more deliberate “reflection-on-action” that takes place after an event (Schön, 1983). It is clear that for some faculty, decisions made quickly and/or without recourse to data are not an uncommon part of their practice.

Yet the literature on decision-making also indicates that expert intuition and rapid decision-making is not always correct, and that certain cognitive biases or heuristics often lead to incorrect responses and decisions no matter how experienced the person. For instance, intuitive heuristics often operate when decision-makers are faced with uncertain situations, and the mind answers easier questions rather than addressing the situation at hand but without noticing the substitution (Kahnemann, 2011). This type of error is particularly common in the case of “System 1 thinking” or rapid, almost unconscious decision-making, as opposed to “System 2 thinking,” or unhurried, deliberate decision-making (Evans, 2003). In addition, Kahnemann (2011, p. 240) cautions that the confidence that a person has in their own intuitions is demonstrably unreliable except in cases where the skill under consideration was acquired in an environment that is regular and predictable (e.g., a chess game), and that opportunities exist “to learn these regularities through prolonged practice.” Given that teaching is a notoriously ill-defined endeavor, it follows that intuitions based on a self-proclaimed expertise in teaching may not necessarily be the most reliable and valid grounds upon which to make decisions.

These perspectives on decision-making suggest two thing in regard to the results reported in this paper. First, while decisions made rapidly in the middle of class are unavoidable and indeed an indispensable and important part of teaching, an exclusive reliance on such System 1 thinking is not desirable. Instead, the slow, deliberate reflection on various forms of evidence after the conclusion of a class or semester is a key component to effective decision-making. As Kahnemann (2011, p. 417) argues, the basis for sound decision-making is the ability to "recognize the signs that you are in a cognitive minefield, slow down, and ask for reinforcement from System 2." Second, some sort of "check" on one's performance as a teacher via a data-based continuous improvement system is a desirable thing, even if one is an expert. This was evident in a biologist who had taught a course over 10 years, yet still carefully scrutinized clicker and exam data in order to track student progress and continually improve his teaching. It is possible that the need for such monitoring may decline as expertise increases, but educators need to carefully balance data and experience as much as administrators and policymakers.

Current Organizational Contexts Are Not Amenable to DDDM

Ultimately, a key finding from this study is that, at the three institutions included in our study, current organizational and departmental structures, cultural norms, and routinized practices are not set up to encourage DDDM. This is evident in the fact that 31 faculty in our study (53%) had developed personalized continuous improvement systems, with no help or assistance from their institutions. This raises the question: Do postsecondary organizations have rules and routines to encourage the careful reflection on data related to student learning and teaching quality? At our study sites we found such self-monitoring at the departmental level to be rare, except in the case of end-of-semester student evaluations, and highlight the following organizational factors that are associated with such a state of affairs.

Few incentives and opportunities exist for faculty to engage in DDDM. 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 or departments where accreditation pressures mandated data collection and reporting, for many faculty the decision whether to collect, analyze, and utilize teaching-related data was left completely up to them. Given that the incentive structure within research universities prioritizes research accomplishments, for many respondents there simply was no compelling reason to commit scarce time to such reflective practices. In any case, while autonomy in one's teaching practice is a hallmark of academic life, the literature suggests that effective DDDM requires not only robust technical systems but also routines and social supports that provide structure to educators' use of data (Mandinach, 2012). Indeed, until cultural norms support regular engagement with data and other forms of evidence to routinely monitor performance and reflect upon the results, it is hard to imagine improvements in organizational decision-making (Kahnemann, 2011).

Little time exists (or is taken) for reflection about data. The notion of reflective practice is central to theories of organizational change and professional development. As Larrivee (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 this reflective process, if it existed at all, entailed brief glances at evaluation data or student exam results largely due to the lack of time available for such activities. Given the demanding workload of most faculty, taking time out of research, teaching, and service activities to engage in reflective practice is simply not tenable. Yet, because reflection is ultimately how "raw" data and information is translated into knowledge, if this stage is missing it is difficult to see how DDDM can be incorporated into academic settings in a meaningful fashion.

Quality data about teaching is unavailable. Within higher education the most commonly available data about teaching itself is the ubiquitous end-of-semester student evaluation. In a study of 72 physics faculty, researchers found that in most cases student evaluations were the only data provided to faculty regarding their teaching, and other sources such as teaching portfolios, peer observations, and research-based assessments that the literature suggests are particularly effective at both capturing the nature of instruction and sparking reflective practice were unavailable (Henderson et al., 2014). In our study, this was also the case and most faculty felt that these evaluations did not provide meaningful data about their teaching for three reasons: (1) poor design, (2) insufficient level of detail, and (3) late delivery of results. The exception to the perceived low utility of evaluation data was students' responses on open-ended questions. In any case, for DDDM to advance in higher education it is clear that institutions should provide additional sources of data about classroom teaching that can complement instructors' own data and insights.

Implications: Insights into Organizational Change and Learning

Using data or "evidence" to improve organizational efficiency and educational quality is becoming a cornerstone of educational policy in the early 21st century. Yet the results from research on DDDM in the K–12 sector is clear—the provision of data alone is not the answer (Coburn & Turner, 2011). However, in the rush towards data mining, learning analytics, and institutional rating systems in higher education, we fear that postsecondary leaders and educators are ignoring the hard lessons learned from data-related initiatives in K–12 schools and districts. For example, Blaich and Wise (2010, p. 67) observe that most postsecondary leaders 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." Instead, as the results reported in this paper indicate, instructional data use implicate not only technical systems but also social networks, cultural norms, and individual dispositions, all of which may vary depending on the context. As a result, for data-based continuous improvement systems to be integrated into standard departmental operations requires those advocating for change to actively understand and then manage the relationships among the potential data users, the system itself, and aspects of the organizational context (Foss, 2014).

While our work represents the beginnings of an evidentiary base for faculty data-related practices that take into account situated practice, future researchers will need to examine faculty data use in additional disciplinary and institutional contexts in order to develop a more comprehensive understanding of these complex dynamics. Furthermore, institutional leaders and policymakers will need to pay careful attention to creating routines, policies, and incentives for instructional data use within academic departments that can set the stage for DDDM. But, in doing so, it will be imperative to recognize that a core feature of successful DDDM is the desire, at the individual and organizational level, to want to continually improve teaching and learning. With this fact in mind, we encourage advocates to take heed of the lessons from the K–12 sector and avoid turning DDDM into a punitive accountability exercise but instead to respect both data and experience.

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

Andrews, T. C., & Lemons, P. P. (2015). It's personal: Biology instructors prioritize personal evidence over empirical evidence in teaching decisions. *CBE-Life Sciences Education, 14*(2), 1–14.

Argyris, C., & Schon, D. (1978). *Organizational learning: A theory of action approach*. Reading, MA: Addision Wesley.

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.

Berger, A. (1997). Continuous improvement and kaizen: standardization and organizational designs. *Journal of Integrated Manufacturing Systems, 8*(2), 110–117.

Bess, J. L., & Dee, J. R. (2008). *Understanding college and university organization: Dynamics of the system* (Vol. 2). Sterling, VA: Stylus.

Bhuiyan, N., & Baghel, A. (2005). An overview of continuous improvement: From the past to the present. *Management Decision, 43*(5), 761–771.

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

Coburn, C. E., & Turner, E. O. (2012). The practice of data use: An introduction. *American Journal of Education, 118*(2), 99–111.

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.

Crandall, B., Klein, G. A., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. Boston, MA: MIT Press.

Darling-Hammond, L. (2010). *The flat world and education: How America's commitment to equity will determine our future*. New York, NY: Teacher's College Press.

Evans, J. S. B. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*(10), 454–459.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative inquiry, 12*(2), 219–245.

Foss, L. H. (2014). Integrating data analytics in higher education organizations: Improving organizational and student success. In Lane, J. (Ed.) *Building a smarter university: Big data, innovation, and analytics* (pp. 187–210). Albany, NY: SUNY Press.

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

Galbraith, J. R. (1977). *Organization design*. Reading, MA: Addison-Wesley.

Gee, J. P. (1986). Orality and literacy: From the savage mind to ways with words. *Tesol Quarterly, 20*(4), 719-746.

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.

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

Gutierrez, K. D., & Rogoff, B. (2003). Cultural ways of learning: Individual traits or repertoires of practice. *Educational researcher*, *32*(5), 19-25.

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.

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.

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.

Henderson, C., Turpen, C., Dancy, M., & Chapman, T. (2014). Assessment of teaching effectiveness: Lack of alignment between instructors, institutions, and research recommendations. *Physical Review Special Topics-Physics Education Research*, *10*(1), 010106.

Hillman, N. W. & Tandberg, D. A., & Gross, J. P. K. (2014). Performance funding in higher education: Do financial incentives impact college completions? *The Journal of Higher Education 85*(6), 826–857.

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*, 1-35. DOI: 10.1002/sce.21175

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.

Huber, G. P. (1991). Organizational learning: The contributing processes and the literatures. *Organization science, 2*(1), 88-115.

Hutchins, E. (1995). *Cognition in the wild.* Boston, MA: MIT press.

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.

Jenkins, D., & Kerrigan, M. R. (2008). *Evidence-based decision making in community colleges: Findings from a survey of faculty and administrator data use at achieving the dream colleges*. New York, NY: Community College Research Center, Teachers College, Columbia University.

Johnston, G. H., & Kristovich, S. A. (2000). Community college alchemists: Turning data into information. *New Directions for Community Colleges*, (109), 63–73.

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY: Farrar, Strauss and Giroux.

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: The Journal of the Human Factors and Ergonomics Society, 50*(3), 456–460.

Kruskal, J, B., & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: Sage Publications.

Lane, J. E. (Ed.). (2014). *Building a smarter university: Big data, innovation, and analytics*. Albany, NY: SUNY Press.

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

Lazer, D., Kennedy, R., King, G., & Vespignani, A. (2014). The parable of Google flu: Traps in big data analysis. *Science, 343*, 1203–1205.

Liou, Y. H., Grigg, J., & Halverson, R. (2014). Leadership and design of data-driven professional networks in schools. IJELM-International *Journal of Educational Leadership and Management, 2*(1), 29–73.

Long, P. & Seimens, G. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review, 46*(5), 30–40.

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

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd Ed.). Thousand Oaks, CA: Sage Publications.

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

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

Picciano, A. G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks, 16*(3), 9–20.

Pesyakhovich, A., & Stephens-Davidowitz, S. (2015). How not to drown in numbers. *The New York Times* (May 3, 2015), 6–7.

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.

Popham, W. J. (1987). The merits of measurement-driven instruction. *Phi Delta Kappan, 68*(9), 679–682.

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

Senge, P. M. (2006). *The fifth discipline: The art and practice of the learning organization*. New York, NY: Random House.

Shavelson, R. J., & Stern, P. (1981). Research on teachers' pedagogical thoughts, judgments, decisions, and behavior. *Review of Educational Research, 51*(4), 455–498.

Simon, H. A., & Gilmartin, K. (1973). A simulation of memory for chess positions. *Cognitive Psychology*, *5*(1), 29–46.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational researcher, 31*(7), 15–21.

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., Halverson, R., & Diamond, J. B. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher, 30*(3), 23–28.

Walsh, J. P., & Ungson, G. R. (1991). Organizational memory. *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.