# Data driven decision-making in the era of accountability: Fostering faculty data cultures for learning


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

The Review of Higher Education, Volume 40, Number 3, Spring 2017, pp. 391-426 (Article)

Published by Johns Hopkins University Press
*DOI: https://doi.org/10.1353/rhe.2017.0013*

*The Review of Higher Education*
Spring 2017, Volume 40, No. 3, pp. 391–426
 (ISSN 0162–5748)

# Data driven decision-making in the era of accountability: Fostering faculty data cultures for learning

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

One of the defining characteristics 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 have 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

---

Matthew T. Hora, PhD is an Assistant Professor of Adult and Higher Education at the University of Wisconsin-Madison where he conducts research on higher education-workforce relations, organizational change, and classroom teaching in postsecondary contexts.

Jana Bouwma-Gearhart, PhD is the Associate Dean of Research in the College of Education, and an Associate Professor of Science and Math Education at Oregon State University, where she focuses on issues related to STEM education, organizational change, and teacher professional development.

Hyoung Joon Park is a Ph.D. student in the Department of Educational Leadership and Policy Analysis at the University of Wisconsin-Madison. His research focuses on utilization-oriented evaluation, evidence-based practices, individual and organizational learning, and STEM education.

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 and then enact corrective measures (Bhuiyan & Baghel, 2005). In education this model has been widely adopted 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 simply a matter of giving educators data reports, but one that involves translating these data into information and actionable knowledge that administrators and teachers can apply to current and future problems (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. Each person will necessarily interpret the data through their own unique perspectives and experiences. Additionally, their situation within a particular school or institution will also influence how they interact with data, including the social networks, cultural norms, artifacts (i.e., designed objects), policies and procedures, and practices that collectively shape how people think and act within complex organizations (Coburn & Turner, 2011; Halverson, Grigg, Prichett, & Thomas, 2007).

Such insights into the processes of sense-making as a situated phenomenon have led to a growing body of research on data use in K-12 contexts known as "practice-based research," which focuses on how educators actually think, make decisions, and work in specific situations rather than on describing the effects of interventions or prescribing best practices (Coburn & Turner, 2012; Little, 2012). In seeking to understand the impacts of the environment on data practices, this line of inquiry emphasizes the cultural aspects of data use, where educators engage in routinized practices with colleagues while using shared language and tools to conduct their work (Spillane, 2012). Given documented challenges with the effective institution of DDDM in schools, particularly at the classroom level, such insights can be an important tool to improve interventions by ensuring that they are aligned with or responsive to the norms and practices of specific organizations, as opposed to a "top-down" approach that is a far less effective approach to reform (Fullan, 2010; Mandinach, 2012; Spillane, Halverson & Diamond, 2001).

What does this all mean for higher education? Policymakers and postsecondary leaders are devoting considerable efforts towards 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 performance-based funding (Hillman, Tandberg & Gross, 2014), institutional rating systems (Kelchen, 2014), and the increasing use of data mining and analytics (Lane, 2014; Picciano, 2012). At the classroom level, some argue that the use of predictive modeling can improve teaching and learning through learning analytics, which is seen as an evidence-based way to tailor instruction to student needs and to generally improve faculty[1] decision-making (Baepler & Murdoch, 2010; Wright, McKay, Hershock, Miller, & Tritz, 2014). Taken together, these developments indicate that higher education has entered an accountability 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, particularly insights gleaned from practice-based research regarding the importance of understanding local data cultures. Besides using such insights to improve the design of new initiatives, they also can shed light on an important question facing the broader field of education – are institutions utilizing technology and data systems to support compliance with accountability pressures or to support learners? (Halverson & Shapiro, 2012). But little is known about how faculty think about and use teaching-related data as part of their regular work and the roles that postsecondary institutions play in supporting the effective use of educational data. This state of affairs is particularly problematic given the tendency for colleges and universities to not engage faculty in continuous improvement systems regarding educational change. As Blaich and Wise (2011) note, the norm is to "gather data, to circulate the resulting reports among a small group of people, and then to just shelve them if nothing horrible jumps out" (p. 12). To successfully take the next step of engaging campus stakeholders—especially faculty—in productive conversations about DDDM will require in-depth knowledge of the faculty cultures for data use.

In this paper we report findings from a practice-based study that examines the cultural practices of data use among 59 science and engineering faculty from three large, public research universities. In this exploratory study we documented how faculty use teaching-related data "in the wild" using interviews and classroom observations, which were analyzed using inductive thematic analysis, exploratory data reduction, and causal network techniques. The study was guided by 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 do these cultural practices and contextual factors play in shaping individual-level practice?

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

## 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 underlying principle is that of continuous improvement, where feedback loops ensure that after an inefficiency is identified, a new method can be tested and results integrated into improved procedures (Bhuiyan & Baghel, 2005). The role that data and other information play within these systems is central, and information-based theories of organizations provide important insights into the relationships between data and change processes, including how individuals perceive, respond, and contribute to knowledge within an organization and collectively affect change (Levitt & March, 1988). In particular, theories such as information-processing (Galbraith, 1977) and organizational learning highlight the fact that data systems are comprised of both technical and socio-cultural elements. This is especially true for an organization's "memory," or the mechanisms whereby data and other information are encoded, stored, and retrieved to inform decisions (Huber, 1991; Walsh & Ungson, 1991). Some researchers on organizational change suggests that when alterations to these memory functions happen, organizational learning has occurred (Levitt & March, 1988), while others claim that transformation to central functions and cultural norms are required for claims of organizational learning (Argyris & Schön, 1978).

### *Data driven decision-making in K-12 contexts*

The idea of continuous improvement spread far beyond matters of business to influence fields as diverse as medicine, 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 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 that had yielded little progress. This legislation and its use of student achievement data as the key component of an accountability system serves as the 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 the key characteristics that comprise organizational data systems. While NCLB privileges *outcome data,* specifically student achievement measured by standardized tests, other data has been utilized including *input data* (e.g., demographics), *process data* (e.g., teaching quality), and *satisfaction data* (e.g., of students and parents) (Ikemoto & Marsh, 2007). Although some have made the case that other forms of information (e.g., homework assignments) should complement test results, the term "data" almost always refers to numeric data in K-12 contexts (Hamilton et al., 2009). Other elements of DDDM include technical infrastructure (e.g., database systems) for managing and reporting data, staff with "pedagogical data literacy" or expertise in using data for educational purposes, time and resources for educators to analyze and interpret data, and leaders who create institutional norms that supports data use and policies that reflect these norms (Hamilton et al., 2009; Liou, Grigg, & Halverson, 2014; Mandinach, 2012).

Of the myriad components that make up K-12 data systems, researchers have singled out two elements as particularly important features: data system design and the social contexts of data use. First, the technological tools available for DDDM such as student information systems, data warehouses, and learning management systems need to be incorporated into a useable and well-designed system of information gathering, analysis, and dissemination (Hamilton et al., 2009). In particular, 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 (Mandinach, 2012). Data practices also involve a social component, as teams of teachers and/or administrators regularly meet to discuss data reports, whereupon the groups develop shared commitments to and views about data use (Coburn, Toure, & Yamashita, 2009). As a result, Spillane (2012) argues that the routinized data-related practices of social groups are the primary grounds upon which interventions succeed or fail. Indeed, some argue that until and unless a "data culture" is developed in which data use is embedded in organizational norms and routines, it is hard to imagine effective DDDM taking place (Hamilton et al., 2009).

Researchers have also documented that administrators and teachers engage in different types of DDDM that range from more to less effective (Marsh et al., 2006). Ikemoto and 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 (e.g., trend data or inferential statistics) forms. Using these criteria, four main types of DDDM are proposed: *basic, analysis-focused, data-focused*, and *inquiry-focused.* The inquiry-based form of DDDM is characterized by the use of complex data analyses as part of a continuous improvement system (Halverson et al., 2007).

Another distinction to be made between different approaches to DDDM centers on the underlying rationale or purpose for collecting and analyzing data – is it to comply with state or federal policies as part of the push for accountability, or is it to support and engage learners? Such a distinction has been made regarding the "culture" of technology use, and it applies to educational data systems as well (Halverson & Shapiro, 2012). Interestingly, when the need for developing human capacity or pedagogical data literacy at the administrator or instructor level becomes evident, then the "learners" are not the students but the adult educators themselves. It is useful, then, to scrutinize an organization's norms and underlying motivations regarding data use (if they exist) to determine whether they primarily support compliance with accountability measures, and/or do they support educators as learners in their roles as instructors and educational decision-makers.

These considerations of culture and social ties are critically important given the simple fact that the implementation of DDDM is not simply a technical issue to be solved by new computers and databases. Instead, the problem is that translating raw data into useable information and actionable knowledge is rather difficult. As Mandinach (2012) argues, "Effective data use requires going beyond the numbers and their statistical properties to make meaning of them" (p. 73). At the heart of this translation process 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). Given that cognitive activity is deeply embedded in the social and organizational contexts in which people operate, factors such as local cultural norms, institutional mission, and organizational routines come into play as an unavoidable aspect of data use (Mandinach, 2012). But whether or not data are noticed by educators in their day-to-day work depends on if the data are perceived as adequately relevant, diagnostic, and valid by educators with respect to immediate problems of practice (Gill, Borden & Hallgren, 2014; Halverson, et al., 2007). Providing such data to educators is much easier said than done, and research indicates that the data are often provided to educators with little attention to utility in the field (Spillane, 2012).

***Data driven decision-making 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 role that data-related systems play in college and university operations. This is changing rapidly as the wave of accountability that engulfed K-12 schools begins to influence postsecondary education, with growing pressure on higher education to embrace a "culture of evidence" (Morest, 2009). While some arguments are rhetorical, others are not unlike the focus on compliance embodied in NCLB, such as performance based funding whereby states allocate funding to public institutions based on data such as student completion rates (Hillman et al., 2014).

At the same time, given the extensive amounts of data available to most colleges and universities, such as graduation rates and tuition revenue, some argue that analytic techniques from the world of Big Data be applied to higher education (Picciano, 2012). For instance, it may be possible to use data mining and related analytic techniques to improve institution-level operations such as targeted recruitment and more efficient admissions operations (Lane, 2014). Additionally, at the classroom level there are now data produced by students via learning management and in-class clicker response systems that effectively represent "learner-produced data trails" that can be analyzed to identify challenging topics or struggling students (Long & Siemens, 2011, p. 32; Wright et al., 2014). The desire to incorporate DDDM into postsecondary classrooms is particularly evident in the science, technology, engineering, and mathematics (STEM) disciplines, where some argue that one of the ways to convince faculty to improve their teaching is to produce evidence or data about the efficacy of certain research-based instructional techniques (Wieman, Perkins, & Gilbert, 2010). Interestingly, while these efforts to improve teaching by providing research evidence and/or results from analytics implicitly require a stage of reflection on the part of faculty, specifics regarding the nature of reflective practice are rarely described, in contrast to the K-12 literature (e.g., Jay & Johnson, 2002).

To date, little empirical research has been conducted on how faculty actually think about and use data as part of their instructional practice. More work exists on challenges facing institutions as they attempt to implement DDDM, such as inadequate capacity in regards to technology and human capital (i.e., skills) for translating data into useful and actionable knowledge (Johnston & Kristovich, 2000). Exceptions that focus on individual actors include research on how biology instructors made teaching-related decisions, which found that personal experience was utilized more than empirical evidence about teaching or student learning (Andrews & Lemons, 2015). In a study of the adoption of data analytics by academic leaders, Foss (2014) found that adoption is shaped by a combination of features, including 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 on the ground in their daily work, results that have been confirmed by other research (Blaich & Wise, 2010; Jenkins & Kerrigan, 2008).

While these studies shed important light on an under-studied topic, the field of higher education would benefit from additional descriptions of data practices at a more finely-grained level. What can be gained by such accounts? Such descriptive analyses provide valuable scientific insights into the current state of affairs regarding one of the most prominent educational reform policies of the early 21st century (Coburn & Turner, 2011), generate new hypotheses and theory for future research (Slavin, 2002), and inform the design of programs and interventions that are aligned with the pre-existing norms and routines of a group of educators as opposed to being completely at odds with local practice (Fullan, 2010; Spillane et al., 2001).

To contribute such descriptive accounts, we build upon the practice-based research tradition to document and describe how faculty use data "in the wild" of their departments and classrooms (Coburn & Turner, 2012; Spillane, 2012). One of the benefits of this perspective is that it avoids the mistake of treating DDDM as simply a technical issue—a "best practice" to be implemented—and instead properly situates it within the realities of the day-to-day pressures and activities of academic work. However, the context is not simply a backdrop to practice, but instead, activity is seen as the relationship among individuals' own attributes, behaviors, tools, and technologies, as well as structural features within a given situation (Greeno, 1998; Hutchins, 1995). These dynamics can see seen in the aggregate as the cultural practices of a group who share similar perceptions of the environment, use similar tools, and engage in similar types of practices (Gutierrez & Rogoff, 2003; Martin, 2003). In this paper we document the cultural practices for data use among 59 faculty and critically examine whether or not their institutions are supporting the effective use of teaching-related data in order to improve educational practice.

## METHODS: ANALYZING CHARACTERISTICS OF DATA USE IN PRACTICE

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

A non-random purposive sampling procedure was used to identify study participants 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, Mechanical Component Design, and Environmental Geology. We contacted

165 instructors via email to request their participation in the study, and 59 ultimately agreed to participate (36% response rate). Participants represented the following disciplines: 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 study site (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%). However, this population was not homogenous as some instructors had long-term contracts while others operated on a year-to-year basis. The course component of interest was the in-class "lecture" and not laboratory or discussion sections.

***Methods of data collection: Eliciting data practices***

A team of four researchers conducted all data collection activities in the Spring semester of 2013. For the interviews we followed the Critical Decision Method (CDM) approach that uses in-depth probes and think-aloud techniques to elicit respondent accounts about a specific recent activity (Crandall, Klein, & Hoffman, 2006). Prior to each interview we reminded respondents that we were interested in a single course that they were currently teaching. 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 in reference to the next class period the respondent would be teaching immediately following the interview. This question was followed by probes regarding the type of data used, specific planning steps, and contextual factors that influenced planning for their next class within a course they were currently teaching (usually within a day or two). The remainder of the protocol included open-ended questions about data-related topics such as continuous improvement efforts within departments. Interviews took place in respondents' offices or nearby conference rooms and lasted approximately 45 minutes.

In addition, each of the participants was observed teaching one or two full class periods using the Teaching Dimensions Observation Protocol (Hora, 2015; Hora & Ferrare, 2013), 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 of the research team members underwent an intensive 28-hour training program over two weeks.

**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 | 12 | 14 | 14 |
| 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 |

***Methods of data analysis: Identifying characteristics of data practices***

First, all interviews were transcribed and entered into NVivo qualitative data analysis software, whereupon two analysts segmented the raw data into smaller units (Gee, 1986). The segments pertained to three core topics that were central to the study: individuals' data routines, the existence of continuous improvement systems, and contextual factors related to data use. Prior to segmenting the entire dataset, the two analysts first applied the topical codes to ten transcripts and then compared coding decisions in order to ensure inter-rater reliability. Next, because the text fragments remained rather complex and lengthy, detailed summaries of each respondent's data use practices were prepared. These summaries distilled the raw data into short descriptions of the three core topics while maintaining respondent language as much as possible. To create these summaries two analysts prepared summaries of ten respondents independently, met to compare results, and made adjustments in order to arrive at a common understanding, whereupon the first author developed summaries for the entire sample.

We then developed a code list comprised of important features of DDDM. The code categories included types of data, types of data analysis, and types of continuous improvement. For each category, we reviewed the literature and included themes 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 includes real-time notes on teaching and direct feedback.

The second step 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). It is important to note that negative or non-responses to types of data use practices were included in this analysis. 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 non-metric 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 six groupings suggested by the analysis in mind, we identified six "types" of cultural practices and the nature of the horizontal and vertical dimensions in the MDS graph.

The third stage of the analysis involved conducting an inductive 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 six types of data use practices were evident at the individual level. The two cases were randomly selected from study participants at 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, as well as other salient factors that may have influenced these decisions. Then we analyzed the observation data for each instructor by calculating the proportions that a particular code was observed in relation to all possible two-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 transcripts to identify whether relationships could be identified based on respondents' explicit statements about associations between any two themes, factors, or behaviors. The results were then used to develop a causal network graphic that depicted the entire decision-making process from planning to classroom instruction (Miles, Huberman, & Saldana, 2014).

Limitations to the study include a small and self-selected sample, the limitation of self-reported interview data in its reliance on respondents' conscious 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 sub-set of disciplines from three research universities, we caution against generalizing the results to broader populations and encourage researchers to use the results to inform future studies with larger, more representative samples.

## RESULTS

### *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 educational context. For example, one biologist expressed confusion at the notion that data would even be used in relation to teaching. 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," we focus on the broad range of data and information resources that faculty utilize in their work. While this perspective departs from the view of data as primarily numeric in the DDDM literature, a broader perspective is consistent with our goal of capturing the types of information considered salient and meaningful to faculty. 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 identified as influencing how the respondent prepared for and monitored their teaching performance. In reporting our findings, we built upon the framework proposed by Ikemoto and Marsh (2007) to categorize the different types of information discussed by respondents in our study (see Table 2).

These data and other information types included those collected prior to the beginning of a course (i.e., input) and those collected during or after a course (i.e., process and outcome). The input category includes various types of data and information that were retrieved in the course of planning, such as colleagues' advice, personal experience, and numeric data. The process and outcome categories refer to those data that faculty gathered and utilized during or immediately after the course such as assessments (both formative and summative) and student satisfaction data (i.e., evaluations provided by institution or instructor). Additionally, two categories refer to information that respondents collected in the classroom regarding student achievement and satisfaction – taking notes in real-time about how well a class went (or not) right after class and also paying attention to direct verbal or written feedback from students.

### *What are characteristics regarding faculty use of data and other information?*

Next, we analyzed descriptions of how faculty actually used these data and information in practice. In doing so, we focused on aspects of data use that the literature suggests are important aspects of data-related practices as well as topics identified in our inductive analysis of the data. These 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 practice (e.g., types of analysis, timing of data use), and for these categories respondent references will not sum to 59 (the sample size). Additionally, in some instances the respondent did not reference a particular aspect of their data practice and thus were not included in the tabulations. Finally, for the 12 faculty who reported no data use at all, calculating their characteristics of data use was not possible. In the interest of space, these instances of non-responses 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 data. The primary distinction between "general" and "sophisticated" types of analyses pertains to the amount of time and degree of detail spent reviewing data. For example, a mechanical engineering instructor who described simply "glancing" at exams and student evaluations, with no evidence regarding in-depth analysis of the data, was categorized as someone

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

| *Data type* | *Example* | *# respondents* |
|---|---|---|
| **Input Data (Prior to course)** | | |
| Personal experience | Memory about student misconceptions, course content, instruction-related issues | 20 (34%) |
| Numeric data | Prior year's exam scores | 11 (19%) |
| Colleagues advice/literature | Insights gleaned from conversations with colleagues and/or research literature | 23 (39%) |
| **Process and Outcome Data** | | |
| *Assessments* | | |
| 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%) |

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

| *Characteristics of data use* | *Example* | *# respondents* |
|---|---|---|
| **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%) |
| **Goal of data use** | | |
| Document student understanding | Goal to document student understanding/misconceptions | 30 (51%) |
| Improve course/curriculum | Goal to apply results to improve course/curriculum | 43 (73%) |
| **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%) |
| **Social nature of data use** | | |
| Solo | Data collection/analysis involves only 1 person | 40 (68%) |
| Group | Data collection/analysis involves 2+ people | 11 (19%) |
| **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%) |
| **Application of data to practice** | | |
| 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%) |

using "general" techniques. This approach was reported by 27 respondents (46% of the study sample). In contrast, analytic techniques considered to be "sophisticated" included 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 to reflect upon student feedback (e.g., office hour conversations, in-class questions) and implications for their teaching.

**Goal of data use/analysis.** An important aspect of inquiry-driven DDDM (Halverson et al., 2007) is the articulation of goals for how data will be used to improve instructional practice. To document this aspect of data use we identified two themes related to faculty goals for data use. First, 30 respondents (51%) discussed using data to "document student understanding" so that they could better diagnose their students' performance. 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." One 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.

**Timing of data use/analysis.** 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 (Coburn & Turner, 2011; Halverson et al., 2007). 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 distinguishing factor among the instructors. For 27 instructors, data were collected, interpreted, and quickly analyzed during class in real time, often leading to an instructional decision in situ. For example, in cases where a large number of students incorrectly answered a clicker question, one biologist reported that she typically changed her lesson plan mid-stream in order to spend more time on the topic. 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 instances, 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. In one case a team-taught introductory biology course involved post-semester meetings where instructors reviewed student assessment data, comments on evaluations, and personal observations about specific teaching activities. The group then made preliminary revisions to the course for the following year.

**Social nature of data use.** Another theme that distinguished one type of data use from another was 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 less effective, we documented the social nature of data use as an important characteristic. 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 together to administer exams and manage data across sections throughout the semester as part of a centralized system.

**Reliance on data expertise.** The importance of staff that are skilled in analyzing and interpreting educational data is well documented, especially in organizations where instructors often lack pedagogical data literacy and/or the time to adequately analyze and reflect on results (Mandinach, 2012). Thus, we recorded the degree to which faculty in the study reported consulting with colleagues or staff who had expertise with educational data. 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 with activities that included assessment design, use of data-related technology, 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 developing concept inventories for the course. In this case, these data were all input into an existing continuous improvement system (i.e., annual program review meetings) focused on 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 analyses of data are more or less effective than ongoing and regular analyses, the latter do suggest that the individual is engaged in a cycle of continuous improvement regarding their curriculum and instruction.

**Application of data to practice.** Finally, one of the critical features of DDDM is that the results of data analyses are actually applied in practice. 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 related to 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.

**Continuous improvement mechanisms.** As previously discussed, at the heart of DDDM is a cycle of continuous improvement where data are 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. 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. 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 influenced the design of continuous improvement systems. Finally, nine faculty reported that external policies (e.g., institutional accreditation) had played a role in creating continuous improvement efforts.

***Do patterns exist in faculty data use practices?***

Next, we examined patterns of data use to identify similar types of practices (i.e., cultural practices). For this analysis we used MDS to explore the similarity among the types of data and information and characteristics of data use practices. The result of the analysis was a MDS graph that depicts the similarities among all interview themes in two-dimensional space, with those themes reported more frequently with one another clustered closely together. The horizontal dimension refers to the degrees of faculty involvement in designing a continuous improvement system from high to low (reading left to right), and the vertical dimension refers to the sophistication of data systems from high to low (reading top to bottom) (Figure 1).

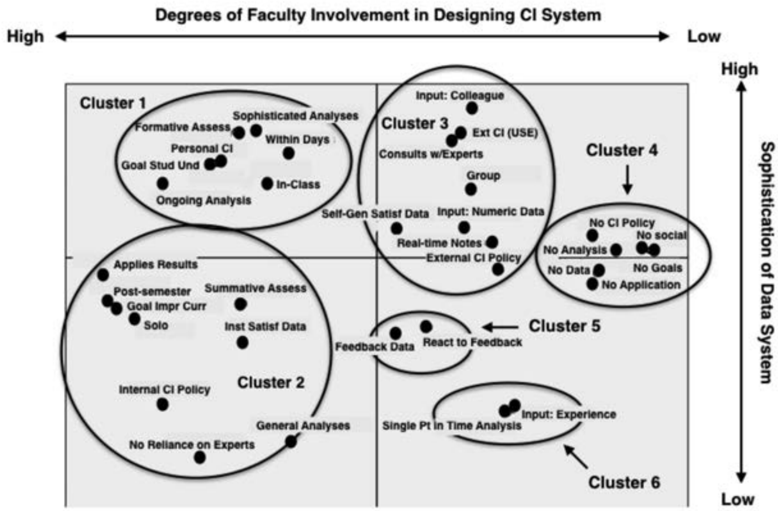


*Figure 1. Multi-dimensional scaling (MDS) analysis of faculty data practices*

The first group (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.

The second group (Cluster 2) contains 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.

The third group (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 type 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).

The fourth group (Cluster 4) includes each of the negative instances 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 from any form of DDDM. The fifth group (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, the sixth group (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 use.

One of the advantages of MDS is that the grouping of variables is not the only story, but the latent dimensionality of the distances must also be interpreted. In other words, the analyst needs to interpret 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 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 continuous improvement efforts were proximal or distal to the educator. This dimension (from left to right) spans efforts that were created by the educator him or herself and/or their own institutions, to those created by external actors, and finally to no policies for continuous improvement at all.

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 can be viewed as more sophisticated than the others, and Clusters 1 and 2 contain continuous improvement efforts designed by educators or their institutions. 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 cultural practices that individual faculty may draw upon in various ways. 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 a single cluster or "type."

***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 were research-related, 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 (-).** One of the constraints facing effective data use is the fact that most faculty lack expertise working with educational data. The skills that respondents reported lacking included the ability to conduct educational research, analyze assessment data to identify patterns and construct implications, manage extensive amounts of data, and to write effective assessments.

**Poor quality of data (-).** Another constraint to effective use of data is the perceived paucity of high-quality data related to teaching provided by their institutions. This complaint focused primarily on one type of data that many respondents felt could in fact help their teaching improve if it were higher quality—end-of-semester student evaluations. 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, which is often months after the end 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 that effectively force administrators and faculty to collect, analyze, and report teaching-related data. This was particularly the case for engineering disciplines, where entities such as the Accreditation Board for Engineering and Technology (ABET) require faculty to collect and report data about student learning in specific competency areas (e.g., the ability to design and conduct experiments) (ABET, 2013). 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 actually reflect on these data and/or find these exercises meaningful or treat them simply as a bureaucratic exercise.

**Policies for course, program, and departmental reviews (+).** In several cases, instructors described formal and informal procedures governing the collection and reporting of data in order to evaluate the quality of courses, degree programs, and entire departments. At the course level, such measures were often required in team-taught courses for which instructors regularly reviewed student assessment data. In other cases, formal program reviews involved the collection and analysis of data (e.g., student exit interviews) in order to assess its efficacy. In both cases, policies for quality assurance essentially dictated the collection and analysis of data.

**Availability of local data experts (+).** Another supportive factor for data use is the existence of other faculty or staff that 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 the hiring of 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.

***Case analyses: The real-world data practices of two biology instructors***

Finally, we examined how individual faculty went about planning and teaching classes in the "natural habitat" of their institutions and departments. For this analysis we used the causal network analysis technique to identify relations among different elements of data practice including perceived constraints and affordances in the organizational context, data practices according to the six clusters identified above, and classroom teaching. The resulting graphs (Figures 2 and 3) depict the inter-connections among these components in order to provide a comprehensive account of data-related cultural practices as they unfold in specific contexts.

**Dr. Robben.** Dr. Robben was a full professor in the biology program at Institution B and was in his 11$^{th}$ 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 data 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."

The causal network graph developed from Dr. Robben's data (see Figure 2) indicated that 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.

**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 (i.e., 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

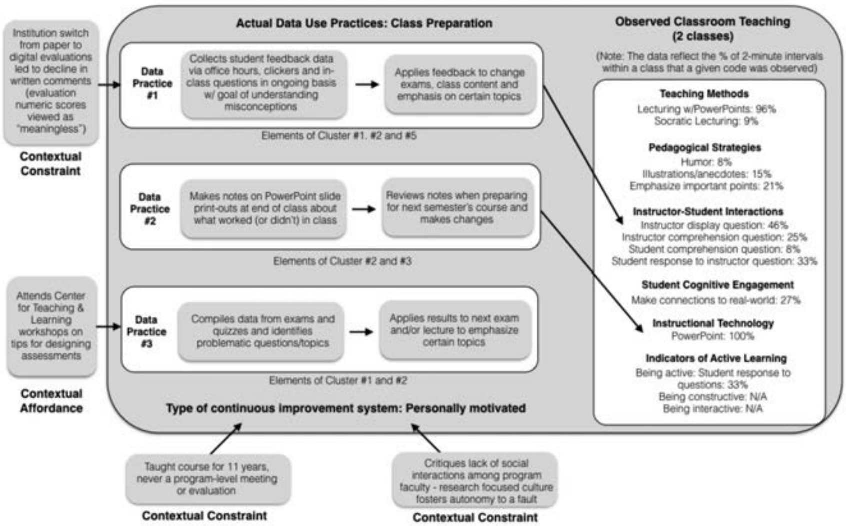


*Figure 2. Data use practices of Dr. Robben*

primary concern with Dr. Iniesta was to identify student misconceptions so that her curriculum and instruction could be improved in the future. 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 learning objectives and assessments, as well as a desire to spend more time with data, she told us that "I would like to do more but I am overwhelmed."

In any case, as is evident in the causal network graph depicting the interrelated components of her decision-making process (see Figure 3), 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 during which she met with the other instructors to review exam data and consider changes to the course. 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%).

## DISCUSSION

This study contributes to the literature on DDDM in general and on data use in higher education in particular in three ways. First, the unique

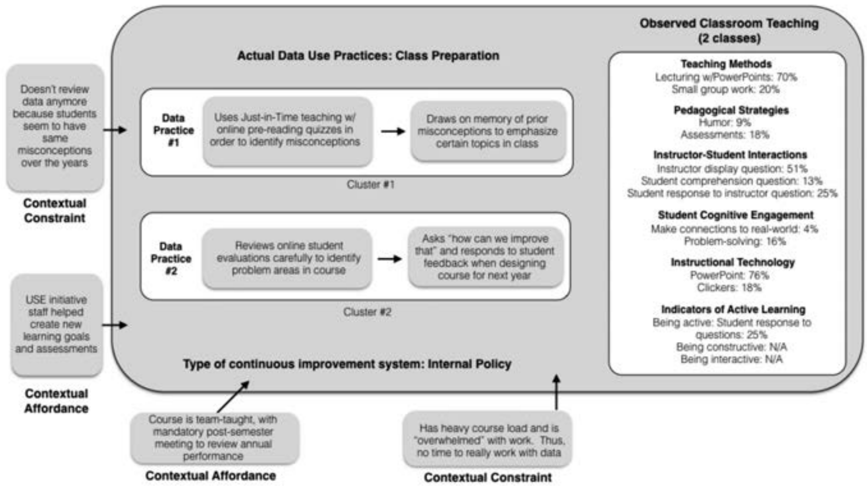


*Figure 3. Data use practices of Dr. Iniesta*

methodology (i.e., using MDS and focusing on cultural practices) utilized in this study represents a new approach to how practice-based research can be conducted in educational settings. While micro-level observations of actual practices are valuable in providing rich and "thick" descriptions of behavior (Little, 2012), it is also important to examine such practices at the meso-level of groups and departments. Multi-dimensional scaling is a technique that is uniquely suited to exploring patterns at group levels, and cultural analyses of these aggregate practices provide an appropriate lens through which to understand these phenomenon (Gutierrez & Rogoff, 2003). Second, the study builds upon prior evidence of DDDM in postsecondary settings by confirming faculty reliance on prior experience and knowledge (Andrews & Lemons, 2015) as well as adding new details regarding the types of data and information used in practice, characteristics of these behaviors, and patterns among these practices. The results also support the finding of Foss (2014) that the salience and utility of data are a critical predictor of use, but the data also indicate that the organizational context plays a considerable role in this process. Finally, by capturing data practices within specific organizational contexts, the results also shed new light on the degree to which these universities are supporting (or not) the effective use of teaching-related data via adequate tools and technology, human resources, and a general culture that supports learners rather than a sole focus on compliance with accountability pressures.

In the remainder of this section we elaborate on this issue of organizational support by first examining details regarding faculty conceptions of data and their real-world processes of decision-making as the grounds upon which any organizational reforms or initiatives must be addressed. Then, by scrutinizing how organizational contexts are supporting or inhibiting these practices, it becomes possible to ascertain whether a culture of accountability or of learning was in place at our three study institutions.

### *Are faculty engaged in DDDM? A closer look at "data" and "decision-making"*

One of the pressing questions facing higher education is whether or not faculty are presently engaged in DDDM in their capacity as classroom teachers. Before we address this question, it is important to consider the fact that most faculty do not receive any formal training in teaching during their graduate training. Thus, for professionals who in many cases are learning "on the fly" how to plan courses, teach in the classroom, and design assessments, it may not be surprising to see less than sophisticated uses of pedagogical data. And our results indicate that for some faculty, the answer to the question about the use of DDDM 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, but more importantly, had no continuous improvement systems in place to inform their teaching. Such an approach is reflected in Clusters #5 and #6 and is also exemplified by the physicist whose idea of quality control was to informally track attendance as a proxy measure for quality.

Our data also indicate that there are many faculty who are engaged in some form of DDDM, but much depends on how you define the term. Some faculty described formal, statistical analyses of numeric data as part of a continuous improvement process that reflects the conventional view of DDDM. 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, mid-term results) to continually update their assessments and lectures. Viewing such practices as a quintessential form of DDDM is implicit in efforts promoting the use of learning analytics and Big Data in higher education.

But there are many other instances where the use of data is less clearly aligned with these views of DDDM. For example, Dr. Robben considered exam results as well as conversations with students and notes taken in real-time to inform his teaching. Given his reliance on not only qualitative data but also ephemeral information (i.e., office hour conversations) as the basis upon which to make his decisions, should his approach be considered data driven? We argue that the answer is yes, but only because there exists evidence of a feedback loop wherein evidence is being carefully considered and then applied to the correction of a situation or problem. Thus, 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 statistical analysis of large numeric datasets 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 components of DDDM – that of "data" and also "decision-making" – constructs that are often left unexamined in the literature. Upon closer scrutiny, it is clear that the distinctions between what constitutes DDDM and what does not are blurry at best.

**The dangers of reifying large-scale quantitative metrics.** In recent years, a backlash of sorts to the Big Data movement has emerged from those 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 humans whose knowledge is essential to contextualize and interpret the results of complex analyses (Lazer, Kennedy, King & Vespignani, 2014). As Peysakhovich and Stephens-Davidowitz (2015) observe, Facebook's data teams are comprised of "social psychologists, anthropologists and sociologists precisely to find what simple measures miss" (p. 6). This point goes back to the finding in K-12 settings that data and sophisticated analytics alone are not the answer to the complex issues facing education, but that people must interpret results and translate them into actionable knowledge (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 have the benefit of conforming to many postsecondary faculty's views of what constitutes valid and reliable data (Wieman et al., 2010). It is simply that these types of data alone do not accurately reflect the breadth of information that faculty find useful and important in their daily work. For instance, while some may argue that practices such as reflecting on the 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. Ultimately, one of the critical points policymakers and educational leaders should consider when designing or implementing data-focused reforms is the fact that faculty utilize a variety information in their daily work.

**Decision-making: The pros and cons of relying on expertise.** Another aspect of DDDM that bears further scrutiny is the nature of decision-making itself. One of the striking findings was the reliance on personal experience as input data and the predominance of rapid, even cursory analyses. In other words, many faculty appear to rely on their intuition and/or expertise. 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. The finding that some faculty rely on "personal evidence," which is often grounded in instructors' perception that they have sufficient expertise to make sound decisions, has also been found by other researchers (Andrews & Lemons, 2015) and termed by education practitioner inquiry scholars as "common sense inquiry" (Boyer, 1990).

It is easy to dismiss decision-making like this as inferior to decisions made with careful analyses. Certainly, it is undesirable for educators to engage in too little to no reflection or consideration about their practice (Schön, 1983). K-12 and postsecondary scholars, alike, have bemoaned teaching as a "private" activity that goes on behind closed doors, rarely examined in any public arena (Shulman, 1993). 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 complex positions and layout of a particular game with a single glance (Simon & Gilmartin, 1973). Evidence also indicates that experienced firefighters are able to make accurate decisions in the midst of crisis through the rapid "search" through memory of particular cues and appropriate responses (Klein, 2008). As Mandinach (2012) suggests, is it possible that an educators' experience alone can lead to a robust knowledge of craft such that the deliberate analysis of spreadsheets are not necessary to inform effective decision-making?

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 decisions. 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 (Kahneman, 2011). This type of error is particularly common in the case of "System 1 thinking," or rapid, unconscious decision-making, as opposed to "System 2 thinking," or unhurried, deliberate decision-making (Evans, 2003). In addition, Kahneman (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). Given that teaching is an ill-defined endeavor, it follows that a sole reliance on one's expertise in the classroom may not be the most reliable grounds upon which to make decisions.

Thus, while decisions made rapidly in the middle of class are unavoidable and indeed an indispensable part of teaching, an exclusive reliance on 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 important. As Kahneman (2011) 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" (p. 417). As with the fact that faculty use a wide variety of data and information in practice, these insights should be carefully considered by those engaged in educational reform. But the prevalence of both System 1 and System 2 forms of decision-making also have implications for what an organization that supports learners would look like.

***Are institutions fostering data cultures for accountability or for learning?***

Recall the distinction made by Halverson and Shapiro (2012) regarding the accountability approach to technology use that is focused on demonstrating compliance with policy, and the learner-centered approach in which technology is used to support the needs and goals of learners. In thinking of the three institutions included in this study in these terms, we arrive at two conclusions. First, that many faculty are engaged in data practices that are focused on helping their students while also helping themselves learn how to become better educators. Second, that current organizational and departmental structures, cultural norms, and routinized practices are not set up to encourage DDDM in ways that support the professional growth of faculty in their capacity as instructors (as opposed to researchers).

**Promising signs of faculty engagement in learner-centered data practices.** The results indicate that many faculty in the study are engaged in data-related practices that reflect a focus on the learner – either themselves or their students. One indicator that led us to this conclusion is the fact that 31 faculty (53%) had developed personalized continuous improvement systems, with no help or assistance from their institutions, because they desired to continually learn how to improve their course and/or their own teaching. It is important to recognize that these cultural practices, however, do not unfold in a vacuum. Instead, elements within the organizational context and even beyond can act to support or inhibit such a focus on learning. For instance, the USE project at two of the study sites was an important supportive influence. By providing post-docs to departments where they acted as local data experts, faculty were able to access in-house resources who could guide efforts to improve the use of pedagogical data as they strove to improve their courses. Such developments should lead to a measured optimism that data cultures for learning can be generated within academia, despite the lack of support from policymakers or institutions.

**Institutions are not supporting data cultures for learning.** These promising examples, however, are unfolding in contexts that are not designed to support data cultures for learning.

**Few incentives and opportunities exist for faculty to engage in DDDM.** Outside of team-taught courses or accreditation policies that 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 the design and implementation of a continuous improvement system.

**Little time exists (or is taken) for reflection about data.** For many faculty, reflection on data, if it existed at all, entailed brief glances at evaluations or student exams largely due to the lack of time available for such activities. Given the 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 realized.

**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 (Henderson, Turpen, Dancy, & Chapman, 2014). Among our study sample, most faculty felt that these evaluations did not provide meaningful data because of their poor design, insufficient detail, and the late delivery of results. 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.

### *Towards fostering data cultures for learning in the era of accountability*

This state of affairs leads to a question – what would a supportive institutional culture look like? First, institutional leaders and policymakers will need 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. Lacking the desire to essentially engage in a form of self-regulated learning will render even the most well designed data system to be a superfluous bureaucratic exercise. Also, until cultural norms that support the regular engagement with data and other forms of evidence to routinely monitor performance are in place, it is hard to imagine improvements in organizational decision-making (Kahneman, 2011). Those caveats aside, we have three concrete recommendations.

**1. Improve student evaluations.** End of course evaluations were roundly denounced by participants in our study as flawed in their quality and the timing of their delivery, which were often months after the completion of a course. While scholars have long debated the pros and cons of student ratings (e.g., Marsh, 1987), it is clear that these data are insufficient for informing DDDM. Thus, institutions should immediately work towards replacing inadequate evaluations with more carefully designed instruments. Additionally, as more institutions utilize online systems response rates must be improved by making evaluations mandatory. Finally, data should be reported back to faculty within weeks of the end of the course, and not months later.

**2. Mandate written reflections on data at end of course.** Institutions should also mandate that faculty spend time at the end of each course to reflect upon data and other information available about the course (e.g., evaluations, assessments) and provide written evidence regarding this process that can then be placed into personnel files. This reflective practice is critical for faculty to "close the loop" of the course planning process whereby final outcome metrics are fed back into a process whereby faculty can think about their implications for the next semester. Given the likely resistance towards any such mandates, leaders will need to be careful to not specify which types of data or continuous improvement systems should be used, but instead respect the autonomy that faculty deserve in running their own courses.

**3. Embed training on DDDM in graduate training.** Finally, many of the behaviors faculty exhibit in their work are vestiges of the training they received as graduate students (Austin, 2002; Oleson & Hora, 2014). In many cases, particularly in the STEM disciplines, this training includes no formal instruction in teaching, much less how to utilize pedagogical data to continually improve one's courses. Efforts such as the Scientific Teaching movement (Connolly, Bouwma-Gearhart & Clifford, 2007; Miller, Pfund, Pribbenow & Handelsman, 2008) that train graduate students in approaching their instruction through the lens of scientific inquiry should be promulgated throughout graduate programs in *all* disciplines.

## Conclusions

While our work represents the beginnings of an evidentiary base for faculty data practices, future researchers will need to examine additional disciplinary and institutional contexts in order to develop a more comprehensive understanding of these behaviors. Other potentially fruitful lines of inquiry include more micro-level research that utilizes observations of data practices in situ (e.g., Little, 2012), examinations of how data-related policies impact (or not) institutions and departments, and critical analyses of how the accountability culture being promulgated across the postsecondary landscape is affecting faculty work and student learning.

Ultimately, 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) observe that most postsecondary leaders assume that the problem of the effective use of data by educators and their organizations is technical and that "once we create sufficiently good measures, widespread institutional improvement in student learning will follow" (p. 67). With this 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 data and experience as the goal of improving undergraduate education is being pursued. As Mandinach (2012, p. 81) argued:

> 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 agree with this sentiment and encourage the field of higher education to shift gears and focus more on creating a culture of data use that supports both faculty and students as they engage in the learning process, rather than a sole focus on using "hard data" to improve institutional efficiency and comply with accountability pressures.

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

Austin, A. E. (2002). Preparing the next generation of faculty: Graduate school as socialization to the academic career. *The Journal of Higher Education, 73*(1), 94–122.

Author (2015).

Author & Author (2013).

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.

Blaich, C. F., & Wise, K. S. (2011). *From gathering to using assessment results: Lessons from the Wabash National Study*. Champaign, IL: National Institute for Learning Outcomes Assessment.

Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.

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.

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

Connolly, M. R., Bouwma-Gearhart, J. L., & Clifford, M. A. (2007). The birth of a notion: The windfalls and pitfalls of tailoring an SoTL-like concept to scientists, mathematicians, and engineers. *Innovative Higher Education, 32*(1), 19–34.

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.

Foss, L. H. (2014). Integrating data analytics in higher education organizations: Improving

organizational and student success. In J. Lane (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: An information processing view. *Organizational Effectiveness Center and School*, *21*, 28–36.

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., & Shapiro, R. B. (2012). Technologies for education and technologies for learners: How information technologies are (and should be) changing schools. WCER Working Paper No. 2012–6. Wisconsin Center for Education Research. University of Wisconsin-Madison.

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), 1–20.

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, 99*(5), 783–818.

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.

Jay, J. K., & Johnson, K. L. (2002). Capturing complexity: A typology of reflective practice for teacher education. *Teaching and teacher education, 18*(1), 73–85.

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.

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

Levitt, B., & March, J. G. (1988). Organizational learning. *Annual review of sociology, 14,* 319–340.

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.

Little, J. W. (2012). Understanding data use practice among teachers: The contribution of micro-process studies. *American Journal of Education, 118*(2), 143–166.

Long, P., & Siemens, 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, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research, 11*(3), 253–388.

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

Martin, J. L. (2003). What Is Field Theory? *American journal of sociology, 109*(1), 1–49.

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

Miller, S., Pfund, C., Pribbenow, C. M., & Handelsman, J. (2008). Scientific teaching in practice. *Science, 322*(5906), 1329–1330.

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. http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

Oleson, A., & Hora, M. T. (2014). Teaching the way they were taught? Revisiting the sources of teaching knowledge and the role of prior experience in shaping faculty teaching practices. *Higher Education, 68*(1), 29–45.

Orlikowski, W. (2000). Using technology and constituting structures. *Organizational Science, 11*(4), 404–428.

Peysakhovich, A., & Stephens-Davidowitz, S. (May 3, 2015). How not to drown in numbers. *The New York Times*, pp. SR6.

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.

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

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.

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

Shulman, L. S. (1993). Forum: teaching as community property: putting an end to pedagogical solitude. *Change: The Magazine of Higher Learning, 25*(6), 6–7.

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.