

Interventions Promoting Educators' Use of Data: Research Insights and Gaps

JULIE A. MARSH

University of Southern California

Background/Context: *In recent years, states, districts, schools, and external partners have recognized the need to proactively foster the use of data to guide educational decision-making and practice. Understanding that data alone will not guarantee use, individuals at all levels have invested in interventions to support better access to, interpretation of, and responses to data of all kinds. Despite the emergence of these efforts, there has been little systematic examination of research on such efforts.*

Purpose/Objective/Research Question/Focus of Study: *This article synthesizes what we currently know about interventions to support educators' use of data—ranging from comprehensive, system-level initiatives, such as reforms sponsored by districts or intermediary organizations, to more narrowly focused interventions, such as a workshop. The article summarizes what is known across studies about the design and implementation of these interventions, their effects at the individual and organizational levels, and the conditions shown to affect implementation and outcomes.*

Research Design: *Literature review.*

Data Collection and Analysis: *This review entailed systematic searches of electronic databases and careful sorting to yield a total of 41 books, peer-reviewed journal articles, and reports. Summaries of each publication were coded to identify the study methods (design, framework, sample, time frame, data collection), intervention design (level of schooling, focal data and data user, leverage points, components), and findings on implementation, effects, and conditions.*

Findings/Results: *The review uncovers a host of common themes regarding implementation, including promising practices (e.g., making data “usable” and “safe,” targeting multiple leverage points) and persistent challenges (e.g., developing support that is generic but also customized, sustaining sufficient support). The review also finds mixed findings and levels of research evidence on effects of interventions, with relatively more evidence on effects on educators' knowledge, skills, and practice than on effects on organizations and student*

achievement. The article also identifies a set of common conditions found to influence intervention implementation and effects, including intervention characteristics (capacity, data properties), broader context (leadership, organizational structure), and individual relationships and characteristics (trust, beliefs and knowledge).

Conclusions/Recommendations: The review finds that the current research base is limited in quantity and quality. It suggests the need for more methodologically rigorous research and greater attention to the organizational and student-level outcomes of interventions, comparative analyses, interventions that help educators move from knowledge to action, and specific ways in which the quality of data and leadership practices shape the effectiveness of interventions.

The K–12 education landscape is lush with data. A significant infusion of federal funds from the American Recovery and Reinvestment Act and Statewide Longitudinal Data System Grant Program has spurred state education agencies to rapidly develop and refine statewide longitudinal data systems that track student test score results, demographic information, and other outcomes. At the district level, administrators have been equally busy collecting their own data and developing systems that include data on students as well as personnel. Education management organizations (EMOs), school reform partners, and nonprofit and for-profit organizations are also contributing to the data onslaught with innovations in assessments and data systems. Within schools, educators have access to an even greater array of data, including formative and diagnostic assessments, teacher-made tests, and student work.

Along with the supply of data, the demand for data use is equally abundant. Policy makers at all levels have clearly articulated expectations that educators use data to drive improvement efforts and track progress (Data Quality Campaign, 2009; U.S. Department of Education, 2009). And although federal and state accountability policies have added pressure and incentives to use data, numerous studies have documented obstacles to such use—most notably, lack of time allocated to educators to examine data and engage in data-related discussions and actions (Feldman & Tung, 2001; Ingram, Louis, & Schroder, 2004); lack of adequate skills and knowledge to formulate questions, select indicators, interpret results, and develop solutions (Choppin, 2002; Feldman & Tung, 2001; Love, 2004; Marsh, Pane, & Hamilton, 2006; Mason, 2002; Massell, 2001; Supovitz & Klein, 2003); and lack of access to timely, high-quality data (Coburn, Honig, & Stein, 2009; Lachat & Smith, 2005; Marsh et al., 2006).

In recent years, states, districts, schools, and external partners have recognized the need to address these barriers and proactively foster use.

Understanding that data alone will not guarantee use, individuals at all levels have invested in *data-support interventions*, or programs, activities, and materials designed to improve educators' access to, interpretation of, and responses to data of all kinds. Interventions range from comprehensive reform initiatives, such as a district that develops a formative assessment system linked to coaching, workshops, and rewards, to more narrowly focused supports, such as a protocol to help educators analyze data or a seminar bringing teams of teachers together to reflect on student work.

Despite the emergence of these efforts, there has been little systematic examination across studies of such efforts to understand their design, implementation, and effects; which components of interventions appear to matter most; and what conditions support or impede success. This literature review seeks to address this gap by synthesizing what we currently know about interventions to support data use and suggesting directions for future research.

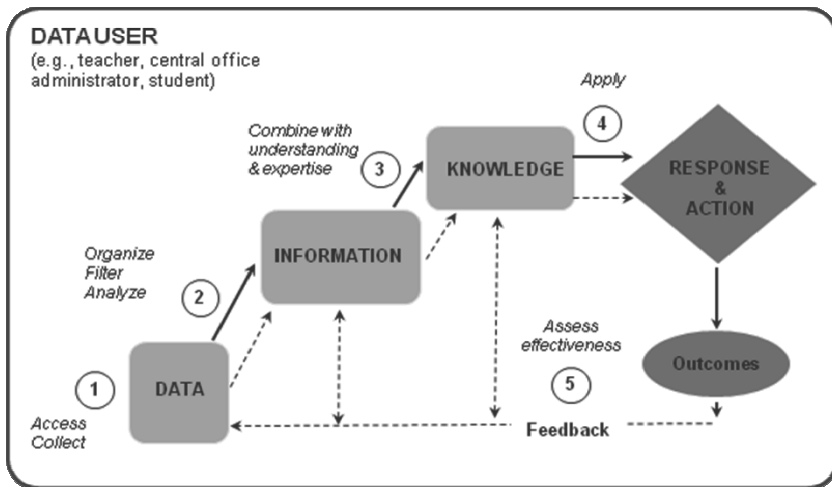
FRAMING THE REVIEW AND METHODS

This review is framed by a theory of action for data use promoted by data advocates and adapted from the literature (see Ackoff, 1989; Mandinach, Honey, Light, & Brunner, 2008; Marsh et al., 2006). First, the definition of data is broad, encompassing not only student test results but also other outcome (dropout and graduation rates), input (student demographic information), process (data on quality of instruction or program implementation), and perception (survey results or opinions from teachers, students, and parents) data. Unlike “evidence” or “information,” the focus here is on raw *data* that must be organized, filtered, and analyzed to become *information*, and then combined with stakeholder understanding and expertise to become *actionable knowledge* (Figure 1). The target “user” could include educators and other stakeholders at the school or district level. Thus, a teacher is expected to apply this knowledge to instructional practice or decisions, whereas administrators and other individuals may apply the knowledge to school- or district-level administrative or managerial practice and decisions (e.g., school or district improvement efforts). The same raw data may point to very different solutions and actions depending on how this process plays out and on the situation and judgment of data users. Once the user has acted and outcomes have resulted, the model assumes that these results, along with new data, can be used to assess the effectiveness of actions, leading to a continuous cycle of collection, organization, and interpretation of data in support of improvement. At any stage, the user may also decide that

more or different kinds of data are required before moving to the next step, leading to feedback loops throughout the process (represented by dotted arrows).

This model implies multiple leverage points at which interventions may occur when supporting data use (illustrated by the numbered circles in Figure 1). An intervention may support users in accessing or collecting data (1), filtering, organizing, and analyzing data into information (2), combining information with expertise and understanding to build knowledge (3), knowing how to respond and taking action or adjusting one’s practice (4), and assessing the effectiveness of these actions or outcomes that result (5). For example, an intervention might include a data system to help teachers access test scores (1), color-coded data displays that organize results by standards and highlight weak and strong areas performance (2), protocols or questions to assist teachers in making sense of these results (3), and training or data coaches who help teachers use this new knowledge to adjust instruction (4) and reflect on these changes and subsequent results (5). Interventions may also target all leverage points (1–5) broadly through a set of expectations, incentives, or norms—for example, a district mandate that teachers participate in data inquiry, or rewards for groups that demonstrate improved outcomes.

Figure 1. Data use theory of action



With this framework in mind, I sought to understand what research tells us about interventions designed to support this process of data use and where there are gaps. More specifically, what are the dimensions or

core features of these interventions? For example, what types of data and data users were targeted, at what leverage point(s), and what strategies were used? How were the interventions implemented, and what were the outcomes? What conditions or factors facilitated or constrained the intervention process?

METHODS

My review process entailed several phases. At first, I searched the electronic databases JSTOR (education only) and ERIC using the search terms “data support,” “supporting data use,” “data capacity,” “data driven,” and “data inquiry.” I also searched the names of researchers known to do work in this area and cited in other articles (e.g., Datnow, Goertz), as well as other articles cited in these pieces, to identify a second set of documents. Finally, I conducted electronic searches for articles or books written about interventions known to the author or cited in other research—including generic terms (e.g., data coach, data team) and specific names (e.g., Bay Area School Reform Collaborative, BASRC). These searches yielded 3,257 documents.

After reviewing all the abstracts, and in many cases, full documents, I carefully selected a subset to include in this review that met the following criteria. I include only pieces that are empirical and not purely advocacy in nature or opinion pieces. To screen for a minimum level of quality, documents come from either a peer-reviewed journal, book, or independent research institution that explicitly submits reports through a review process. I include only pieces that focus on efforts to support data use at the K–12 level and with enough detail or substance to probe key areas I outline next. For example, many documents found in my searches and excluded from this review focus primarily on how educators use data and may have described conditions that support data use briefly at the end. Other documents included in the review may have focused on data use more broadly or a reform effort that involved more than support for data use, but devoted significant attention to the data-related interventions and were therefore included in the review. Per the mentioned specifications for defining data-support interventions, I include studies that define broadly both data and data user. As such, studies that describe their focus as “evidence” but include “data” as one aspect of evidence were included, as were studies of educators or administrators at the district or school levels.

To further narrow the scope, I exclude research on interventions involving only one school or classroom and focus on studies of broader system-level (e.g., local education agency, school reform network) or

multischool interventions (e.g., workshop for school teams). I also focus on interventions targeting general education, excluding a large body of special education research that deserves a separate review. Finally, although some of the interventions reviewed in these studies include technology-related components as well as formative assessment tools, I do not include in this review studies that *exclusively* focus on technology support because this research field was recently reviewed by Halverson and Shapiro (in press), or the use of formative assessments, recently reviewed by Young and Kim (2010) and Supovitz (2012, this issue).

Ultimately this careful sorting process yielded a total of 41 books, peer-reviewed journal articles, and reports for this review, listed in the appendix. After reading each document, I wrote and coded a summary to identify the study methods (design, framework, sample, time frame, data collection), intervention design (level of schooling, focal data and data user, leverage points, components), and findings on implementation, effects, and constraining and enabling conditions.

In some cases, several of these pieces examine the same intervention. Either different researchers studied the same intervention, or the same researchers published findings in multiple sources using a different theoretical lens or focusing on a different facet of the intervention. When relevant, I refer to these pieces as a group, as collective research on that intervention. Depending on the context, when reporting frequencies of findings throughout this review, I refer to the denominator as either the number of pieces reviewed (41) or grouped studies/interventions (29). Finally, although the majority of these studies were conducted in the United States, I include five international studies because they focus on types of data (two focus on value-added test results, two on high-stakes tests and classroom observations, one on student engagement data) and interventions activities (e.g., workshops, data teams) commonly found in U.S. schools.

The remainder of this article is organized in five sections. First, I comment on the studies reviewed and methods employed, followed by a brief description of the interventions examined. Next, I synthesize key findings about implementation across these studies. I then summarize what the research tells us about outcomes of these interventions at the individual and organizational levels, as well as the conditions shown to affect implementation and outcomes. I conclude by presenting an expanded framework for examining interventions in the future and suggestions for where more research is needed.

OVERVIEW OF STUDIES AND INTERVENTIONS

The reviewed documents generally cluster into three broad groups. The first two groups of studies examine initiatives sponsored by local school systems (districts, education management organizations, local education agencies) and intermediary organizations (school reform model sponsors/networks, university partners, nonprofits). Most of these system-level initiatives tend to be comprehensive in their attempts to support data use with multiple strategies and leverage points (see details that follow). Two thirds of them focus on one case, such as Duvall County Public Schools (Supovitz, 2006; Supovitz & Weathers, 2004) or the University of Pittsburgh's Institute for Learning (IFL; Honig & Ikemoto, 2008; Kerr, Marsh, Ikemoto, Darilek, & Barney, 2006).¹ The other third look across multiple districts or intermediary organizations, such as the University of Southern California (USC) research on four high-performing districts (Park & Datnow, 2009; Wohlstetter, Datnow, & Park, 2008) and the Consortium for Policy Research in Education's studies of 23 districts (Massell, 2001; Massell & Goertz, 2002). The third set of studies reviewed pertains to professional development initiatives involving more than one school (state coaching initiative, workshop for teams from multiple schools). Some of these initiatives involve small numbers of participants, and their studies tend to examine microlevel processes.²

As illustrated in the appendix, the overwhelming majority of studies reviewed are qualitative in nature, employing case study designs, interviews, focus groups, observations, reviews of documents, and, in some cases, cross-sectional surveys of data users. Only seven pieces included comparison groups in their designs—including four that employed quasi-experimental designs and one randomized controlled trial.³ More than a third of the pieces (15) relied on 1 year of data, and only seven drew on more than 4 years of data.

Across all three categories, the studies exhibit varied levels of methodological rigor.⁴ Some authors provide strong assurances of trustworthy and credible findings,⁵ including sufficient detail about data collection and analysis methods, triangulating data sources and/or methods, and providing rich descriptions and ample supporting evidence (some of the most notable are Gearhart & Osmundson, 2009; Goertz, Olah, & Riggan, 2009; and Means, Padilla, & Gallagher, 2010). Other studies, however, are more mixed in the level of quality, failing to provide sufficient detail about methods (e.g., Ancess, Barnett, & Allen, 2007), to triangulate (e.g., Denton, Swanson, & Mathes, 2007), and/or to provide sufficient evidence to back up findings (e.g., Yang, Goldstein, Rath, & Hill, 1999). Other studies did not attain high response rates on surveys (e.g., Kerr et

al., 2006; Quint, Sepanik, & Smith, 2008) or failed to provide information about sample sizes or response rates (e.g., Demie, 2003). Many studies employing quantitative methods did not provide details on reliability and validity of instruments and measures, and only a handful of qualitative studies included respondent validation or “member checks.” In nine pieces reviewed, one or more authors were participant observers involved in designing and/or delivering the interventions. In some of these cases, these authors provided assurances of credibility through member checks and triangulation (e.g., Slavit & Nelson, 2010), but in other cases, these assurances were lacking (e.g., Murnane, Sharkey, & Boudett, 2005). Finally, only 7 of the 41 reviewed pieces apply theoretical frameworks to the analyses (e.g., principal-agent theory in Wohlstetter et al., 2008, and sociocultural learning theory in Honig & Ikemoto, 2008), and another 15 relied on either the theory of action of the intervention (e.g., Gallimore, Ermeling, Saunders, & Goldenberg, 2009) or a conceptual framework derived from literature (Anderson, Leithwood, & Strauss, 2010) to guide analysis. The remaining 19 articles did not present an explicit framework.⁶

These strengths and weakness are important to keep in mind when considering the findings presented in the remainder of this article. I will return to these methodological issues throughout the review and have included details on all studies in the appendix so that readers can judge for themselves the strength or credibility of findings of particular authors.

CORE DIMENSIONS OF THE INTERVENTIONS

Domains of Action

The interventions examined in these studies adopted many strategies to deliver this support, which fell into five key domains, outlined with examples in Table 1: human support, technology support, data production, accountability and incentives, and norms and expectations. As the appendix indicates, all the interventions reviewed involved some type of human support, particularly professional development. About a third included technology support (10 of the 29) and data production (9 of 29), and even fewer explicitly included accountability mechanisms (4) and norms and expectations (6). A recent nationally representative district survey sheds light on the prevalence of some of these strategies:

- Human supports: A total of 80% of districts provide schools with technical experts to help access data from data systems; 50% provide

at least some schools with access to “data coaches”; 65% provide teachers with processes or tools to effectively utilize data for instructional purposes; and more than 90% provide at least some school

Table 1. Data Use Intervention Domains of Action: Examples From the Literature

	Examples
Human Support	
<i>Professional development</i>	<ul style="list-style-type: none"> • Global Perspective for Local Action includes a 2-day workshop for district-level teams to collaboratively examine TIMSS data and follow-up support for using action research to generate local action (Huffman & Kalnin, 2003). • Success for All trainers must conduct a case study of a school, analyzing data on successes and challenges, and review these case studies with colleagues at annual conferences. (Park & Datnow, 2008).
<i>Coaching</i>	<ul style="list-style-type: none"> • One key responsibility of site-based literacy coaches in Florida is to support teachers’ use of reading diagnostic and progress monitoring data (Roehrig et al., 2008; Marsh et al., 2010). • Boston’s FAST-R initiative employs data coaches who support participating schools with the analysis and use of the interim assessment results (Quint, Sepanik & Smith, 2008).
<i>Tools</i>	<ul style="list-style-type: none"> • A comprehensive school reform model requires the use of an “inquiry protocol” outlining a process for analyzing student work, identifying shared student academic problems, and developing and planning instructional responses (Gallimore, Ermeling, Saunders & Goldenberg, 2009). • The Institute for Learning’s Learning Walk protocol guides educators through the process of collecting and analyzing data about instructional quality (Kerr et al., 2006; Honig & Ikemoto, 2008).
<i>Technical expertise</i>	<ul style="list-style-type: none"> • One UK LEA research office translates complex raw and value-added test data into user-friendly tables, graphs, and reports for schools (Demie, 2003). • Intermediary organizations help the Oakland Unified School District manage and simplify complex data and information (Honig, 2004).
<i>Networking/ brokering</i>	<ul style="list-style-type: none"> • A research center’s “liaison” link’s up school data teams with external resources and supports to address needs identified by data analysis (Bond et al., 2001).
Technology Support	<ul style="list-style-type: none"> • Milwaukee Public Schools supports four technology systems and tools that make data accessible to schools; some generate customized reports and others allow users to engage in open-ended “data discovery” (Moody & Dede, 2008).
Data Production	<ul style="list-style-type: none"> • To fill a void left by state testing, the Merck Institute for Science Education Partnership helped develop various science performance assessments and tasks for use in partner districts (Corcoran & Lawrence, 2003).
Accountability & Incentives	Edison’s “star rating system” recognizes and rewards schools and educators for ratings on measures of customer satisfaction, operational excellence, student achievement, and implementation of school design which includes the use of interim assessments (Marsh et al., 2008).
Expectations & Norms	States and districts require or encourage school improvement planning or site-based decision-making with an expectation that data are central to planning (Massell, 2001).

staff with training designed to enhance school capacity to use data to improve instruction.

- **Technology:** More than three-quarters of districts have data systems that organize and analyze interim assessment data and data warehouses with current and historical student data.
- **Formal expectations:** A total of 69% of districts require all or some schools to follow specific data-driven decision-making practices when developing their school improvement plans, and 65% articulate for teachers specific processes for how to use data to guide instruction (Means et al., 2010).

Not surprisingly, researchers also found that larger districts had more supports than smaller ones.

Types of Data

All but three of the interventions studied focused at least in part on supporting the use of student performance data from state exams; value-added, interim, and diagnostic assessments; transcripts/grades; or student work (often along with input data such as student demographics). Among these interventions, more than three fourths focused on performance data in English language arts (ELA) and/or mathematics (21), and almost half included science data (12). Seven interventions examined other student outcomes, such as data about student engagement collected via student surveys (Bond, Glover, Godfrey, Butler, & Patton, 2001) and classroom observations (Nelson & Slavit, 2007). Ten interventions supported the use of perception data, such as Milwaukee's climate surveys (Moody & Dede, 2008) and the Gatehouse Project's student surveys of learning environment (Bond et al., 2001), and 11 involved process data, including peer observations and videos of classroom instruction (e.g., Nelson & Slavit, 2007; Stringfield, Reynolds, & Schaffer, 2008) and assessments of reform or design implementation (e.g., Marsh, Hamilton, & Gill, 2008; Supovitz, 2006). Seven others supported the use of research in addition to data.

Data Users

The majority of interventions targeted support for teachers (26 of 29) and/or principals (20), and a third (10) included central office administrators as focal data users. Only four interventions conceived of data users more broadly to also include community members and students (Bond et al., 2001; Huffman & Kalnin, 2003; Marsh, 2007), counselors

(Bond et al., 2001), and school board members (Huffman & Kalnin, 2003; Marsh, 2007). More than half of the interventions focused on K–12 or K–8 schooling (17), and the remaining interventions focused only on educators at either the elementary (9) or secondary (3) level.

Leverage Points

Although two interventions exclusively supported the front end of the data use process (leverage points 1–3)—the U.K. Local Education Agency (LEA) efforts to help educators interpret raw and value-added test results (Demie, 2003; Yang et al., 1999)—the other interventions included efforts to support users at multiple leverage points (1–5). For example, the four exemplary districts studied by USC researchers attempted to help educators access (1), organize, and interpret data (2), build skills and knowledge needed to apply what is learned to practice (4), and monitor results (5). Similarly, professional development and facilitated inquiry team meetings examined in one study attempted to help math and science teachers with framing questions, collecting and interpreting data (1–2), and reflecting on and adjusting instructional knowledge and practice (3–4; Nelson & Slavit, 2007).

IMPLEMENTATION THEMES

Several common themes emerged across these studies with regard to implementing data support interventions. These themes explore what happens when well-intentioned individuals and organizations try to enact interventions and examine the *process* of helping educators gather, organize, interpret, and act on data. The first set of themes focuses on promising practices and features of interventions, and the second set highlights common tensions and challenges to implementation.

PROMISING PRACTICES AND FEATURES

The first set of themes identifies practices or features associated with “successful” implementation. In most cases, success is defined as interventions being implemented as designed and the process running smoothly (e.g., educators regularly met, participated in the intervention activity, or used the tool). In a minority of studies, authors linked these promising practices to outcomes—either a general sense that “effective use of data” ensued or, less often, that specific changes in school or individual practice resulted (e.g., development of a learning community, new instructional practices). As such, the themes should be interpreted with caution.

In many cases, there is not consistent, rigorous evidence proving that these practices lead to data use or other desired outcomes but instead a general sense that they contribute to better implemented interventions. I attempt to clarify the meaning of success and the evidence provided throughout this discussion.

Making Data “Usable” and “Safe” Are Important Preconditions For Use

Six of the 41 studies indicate that translating data into comprehensible and simplified forms is an essential first step in supporting data use (Demie, 2003; Honig, 2004; Kerr et al., 2006; Means et al., 2010; Murnane et al., 2005; Stringfield et al., 2008). For example, researchers analyzing the implementation of the High Reliability Schools (HRS) model identified as a key “theme” the importance of developing methods for disaggregating test results and the centrality of software that allows educators to do so efficiently (Stringfield et al., 2008). They did not, however, provide evidence linking these practices with outcomes. Another study found that two districts in which educators reported greater frequency of data use relative to a third district purposefully assigned individuals to “filter” data—employing either staff in the central office with strong data-analysis skills to make data more usable for school staff (e.g., by completing initial analyses and summarizing results in easy-to-understand tables and graphs) or school-based coaches who often took the first step of analyzing test results and presenting them in usable forms to their colleagues on site (Kerr et al., 2006). Although this research associates filtering with greater use, it does not substantiate a causal relationship.

Making data “safe” appears to be another prerequisite for facilitating data use. Once again, however, the research base supporting this finding is mixed in its methodological rigor. Several studies from the United Kingdom related the success of the LEA in helping schools understand and use test score data to their commitment to ensure confidentiality and their consistent message that the data are “a tool to raise questions rather than make judgments about the school” (Demie, 2003, p.42; also Yang et al., 1999). The link to “success” and actual data use in both studies, however, is tenuous and rests on research with limitations noted previously (i.e., self-reports, no documentation of response rates). Nevertheless, this notion of safety is a consistent theme highlighted in other higher quality studies. In case studies of high-data-using schools, school leaders reported that it was important to start with nonthreatening uses of data, such as anonymous classroom-level test results, to build mutual trust among teachers participating in data inquiry groups (Means et al., 2010).

Another district study, however, suggests that assuring anonymity may not be sufficient for gaining the trust and cooperation of participants (Supovitz & Weathers, 2004). Despite assurances that data collected via school walkthroughs would remain anonymous, some educators remained guarded throughout the process, fearing that district leaders would use the data for evaluative purposes.

Still other studies identified particular structures, practices, and rules employed in ways that facilitated safe and productive discussions around data and provided evidence supporting these connections. For example, USC researchers found that in all four data-driven districts, principals frequently set formal expectations and rules governing meetings around student data to ensure that conversations were productive, purposeful, and not personal in nature (Park & Datnow, 2009). In another study, intermediary organization staff helped principals focus on the data—in this case, videos of classroom instruction—by setting up norms and explicitly encouraging principals to focus on the evidence of quality instruction instead of evaluating the individual teacher. Authors explained that specific “talk moves” utilized by intermediary staff helped structure conversations and limited the “risks” of collectively analyzing and critiquing instructional practice (Honig & Ikemoto, 2008). Similarly, facilitators of inquiry teams in another intervention helped teachers use structured protocols to create “safe spaces” for data analysis and prevent “wandering” discussions (Nelson & Slavit, 2007).

A few studies found that formal rules and protocols governing data discussions not only help promote safety but also guard against particular individuals dominating conversations. For example, facilitators of strategic planning sessions in one district enforced a norm of “leave your title at the door,” which ensured that decisions around what reforms the district should implement were driven by careful deliberation around data and ideas instead of the power of any particular individual. Contrasting sharply with another district that did not use these strategies and in which observational data showed district leaders controlling conversations and decisions, participants in the district that adopted these norms reported feeling free to state their concerns, believed that participation was fairly evenly distributed, and cited examples in which ideas with wide appeal rose above the positional power of participants. A climate of trust, however, also contributed to the acceptance of these norms and the differences observed across sites (Marsh, 2007). Another study provides limited evidence that structured protocols and tools were valuable for promoting involvement in data discussions (Murnane et al., 2005).

Interventions May Have Greater Traction When They Are Comprehensive and Target Multiple Leverage Points

Several researchers concluded that one support alone does not lead to changes in behavior or actual use of data. These studies found that just providing a data system to improve access (leverage point 1; e.g., Goertz et al., 2009), a training session to assist with translating data to information, actionable knowledge, and action (2–4; Gallimore et al., 2009), or both (Means et al., 2010), will not lead to data use. In one quasi-experimental study, interveners recognized that schools were not implementing grade-level data teams after receiving brief training and augmented the intervention to include ongoing training, technical assistance, and a formal protocol codifying the inquiry process to help teams translate information into actionable knowledge. After making these adjustments, researchers found consistent evidence of differences between the frequency and nature of data team meetings in treatment and comparison schools (e.g., treatment meetings were more focused on purposeful use of assessment data and efforts to implement and evaluate jointly developed instructional strategies; Gallimore et al., 2009). Other studies similarly emphasized the importance of linking tools with in-depth professional development to assist users with understanding the purpose of and theory behind the tool and how to apply it in practice (Gearhart & Osmundson, 2009; Kerr et al., 2006; Supovitz & Weathers, 2004). Nevertheless, even with focused support, users do not always apply the tool as intended, as illustrated by the superficial implementation of the IFL Learning Walk tool documented by Honig and Ikemoto (2008) and Kerr et al. (2006). Others have similarly noted that participants' prior knowledge, beliefs, and attitudes greatly influence the way they interpret and implement interventions (Coburn, 2010; see also Spillane & Miele, 2007, not part of this review: I return to this topic later under "Conditions").

Still other studies indicate a need for even more comprehensive supports. For example, to achieve greater consistency in the level of data use districtwide, one high-data-use district profiled in Means et al. (2010) added to its initial data system expectations and training for schools to establish professional learning communities (PLCs), along with systems of monitoring and rating the quality of PLCs and the frequency of teachers' use of data systems.. This study and others also highlight the importance of targeting multiple leverage points of data use. For example, USC researchers concluded that a constellation of strategies supported data use in the four systems they studied, including investing in data management systems and identifying data appropriate for informing educators'

work (leverage points 1–2); building capacity to interpret and act on data through professional development, facilitation, time, networking, and tools (2–4); and defining goals and establishing a culture of data use (1–5; Park & Datnow, 2009).

Collaboration—Horizontal and Vertical—Appears to Be an Important Component of Successfully Implemented Interventions

Sixteen studies identified opportunities for cross-department, school, and, in some cases, district interactions as key ingredients of support for data use (Anderson et al., 2010; Gearhart & Osmundson, 2009; Goertz et al., 2009; Honig & Ikemoto, 2008; Huffman & Kalnin, 2003; Kerr et al., 2006; Marsh et al., 2008; Massell, 2001; Murnane et al., 2005; Nelson & Slavit, 2007; Park & Datnow, 2009; Porter & Snipes, 2006; Stringfield et al., 2008; Supovitz, 2006; Supovitz & Weathers, 2004; Wohlstetter et al., 2008). Some of these studies provide concrete evidence of links between opportunities to collaborate and desired outcomes. For example, in one study of 43 districts, researchers found that central office administrators played a significant role in facilitating principals' use of data, which included opportunities to collectively share and interpret data with principals from other schools in monthly meetings (i.e., vertical collaboration). Survey data indicated significant correlations between reports of district support—which included collaborative opportunities—and reports of using evaluation and student achievement data to help make decisions (Anderson et al., 2010).

In Duvall County, cross-school visitations in which educators collect “snapshot” data on the quality and depth of reform implementation was said to be a catalyst for learning on the part of both the visitors and those visited (Supovitz, 2006; Supovitz & Weathers, 2004). Time set aside in monthly principal meetings and central office administrators' purposeful facilitation of collective analysis of the snapshot results contributed to this process. As one central office administrator explained,

I think it is especially valuable to hear conversations between principals from different regions about what they saw and what their understanding was and what their training has been and how they have accomplished things, especially with moving their teachers, because I think really the most difficult part is not the presentation of information, it's having staff to adopt it, and how do you get that to occur. (Supovitz, 2006, p. 183)

Similarly, the most common feedback given to practitioner researchers

studying a professional development effort involving teachers in collective examination of student data and instructional practice was that participants “valued the opportunity to have focused conversations with other teachers in the same building but in a different department (mathematics and science) and with teachers of the same subject at other grade levels” (Nelson & Slavit, 2007, p. 28). Other researchers similarly found that participants appreciated cross-site collaborative opportunities (Huffman & Kalnin, 2003; Murnane et al., 2005; Stringfield et al., 2008). In one evaluation of the comprehensive school reform model High Reliability Schools (HRSs), teachers frequently cited as valuable learning experiences observing classrooms in higher achieving, demographically similar schools and off-site, cross-school retreats examining best practices (Stringfield et al., 2008).

A few studies also found that involvement of different stakeholder groups or individuals with different roles in the system contributes to the collective interpretation and use of data (Bond et al., 2001; Marsh, 2007). For example, research on a whole-school adolescent health intervention found that including in data teams representatives from the entire school community—including administrators, teachers, parents, students, curriculum leaders, counselors, and student welfare coordinators—to examine survey data on student perceptions of the school’s social and learning environment ensured that responses were owned and enacted by the entire school (Bond et al., 2001).

IMPLEMENTATION TENSIONS AND CHALLENGES

Another set of themes that emerged across studies pertains to challenges and tensions that interveners faced when implementing data support. These findings generally describe difficulties encountered when trying to support educators in all stages of the data use process in complex educational environments—contexts in which data users have competing needs, multiple data-support interventions coexist, and resources and capacity are limited.

Interventions Face a Tension Between the Provision of Tailored Versus Generic and Concrete Versus Conceptual Support to Data Users

Six studies suggest that one challenge in implementing interventions is deciding the level of specificity and customization of data supports. This theme is particularly relevant to studies of intermediary organizations and school reform networks.⁷ For example, in the early years, BASRC started work with districts and schools focused on standardized tools and

strategies and in later years developed customized support using coaches and more tailored tools in response to negative feedback. Yet even after making these changes, BASRC failed to truly meet diverse needs of its partner sites. In some settings, teachers felt that the protocols did not relate to their school or students; in other cases, teachers felt that the protocols were too generic or basic. Throughout its history, BASRC struggled to achieve a balance between “prescribing and co-constructing” the reform efforts with partner sites, which proved to be very difficult given the diversity of partner schools and districts it served (Jaquith & McLaughlin, 2009). Researchers evaluating HRSs attribute success (in this case, student achievement gains) to the project’s emphasis on “on-the-ground co-construction of school reform” rather than strict fidelity of implementation of the model’s original set of ideas. Maintaining a common set of general principles and allowing for local adaptation “filled in details as to how to successfully implement abstract academic concepts and greatly increased local ownership of the process” (Stringfield et al., 2008, p. 425).

Multiple studies of the IFL identified a dilemma of district partners appreciating a lack of specificity in the work of the IFL because it allowed for district ownership and adaptation, and simultaneously requesting more practical advice on how to operationalize IFL ideas, which were often viewed as overly theoretical (Honig & Ikemoto, 2008; Kerr et al., 2006). IFL staff in turn faced a dilemma of wanting to develop concrete tools to help implement ideas but fearing that these efforts might “proceduralize” and lead to inauthentic translation of IFL ideas and theory (Kerr et al., 2006). In another study of a portfolio assessment tool intended to help teachers interpret student work to guide instruction, researchers concluded that although the generic protocol contributed to valuable learning and development of a PLC, more targeted support was needed to address teachers’ particular curricular areas and personal goals (Gearhart & Osmundson, 2009). Leaders and evaluators of the school reform network Getting Results similarly reported learning over time that although the model was curriculum free and process oriented, there was a need to provide concrete support strategies (e.g., protocols for analyzing student work in groups, detailed procedures on how to administer, score, and analyze assessments). They concluded that it is “valuable to try and work on both planes with school staffs—on the concepts or principles underlying concrete procedures and the successful execution of those procedures” and that providing both types of support contributed to higher levels of model implementation over time (McDougall, Saunders, & Goldenberg, 2007, p. 85).

Studies Yield Mixed Findings About the Role of Accountability Mechanisms in Promoting Data Use, in Some Cases Leading to a Tension Between Providing Pressure Versus Support

Some researchers have indicated that accountability mechanisms and incentives are important aspects of interventions to promote data use. These elements refer not to broad accountability policies such as the No Child Left Behind Act but instead to formal and informal means of motivating and holding educators responsible for participating in and following through with intended purposes of interventions. For example, USC researchers noted that in addition to state and federal accountability policies, school system requirements to develop school improvement plans and policies linking compensation to student performance “motivated schools to carry out the data-driven activities the systems requested” (Wohlstetter et al., 2008, p. 249). One quasi-experimental study that found greater use of data in faculty meetings and shifts in teacher thinking in treatment schools employing inquiry protocols concluded that “these outcomes are unlikely in the absence of building leadership that supports and holds teacher teams accountable for sustaining the inquiry process until they see tangible results” (Gallimore et al., 2009, p. 544). According to other researchers, a lack of accountability in the form of clear expectations and measurable goals contributed to the uneven implementation of data use practices across BASRC sites (Jaquith & McLaughlin, 2009). Most of these studies, however, do not provide strong evidence supporting these assertions.

In contrast, other studies find unintended, negative consequences of accountability mechanisms and incentives on data use. For example, a study of multiple interventions in Milwaukee Public Schools demonstrated that a lack of coherence among the accountability purposes and design of interventions led to vastly different responses from data users (Moody & Dede, 2008). For example, one set of technology tools designed primarily for external accountability purposes generated lists of students close to state proficiency targets (“bubble kids”), which contrasted with another set of interactive and flexible tools designed for internal accountability purposes that promoted open-ended exploration of patterns. Whereas the former led teachers to instructional practices targeting particular students who could elevate school-level performance measures, the latter promoted more reflective instructional practice and discussions among colleagues. Researchers concluded that such differences led to “a fracturing of the district’s efforts to support this type of work” (Moody & Dede, 2008, p. 234).

Some studies also make an important distinction between formal and

informal or external and internal accountability mechanisms, suggesting in some cases that norms and pressures from colleagues may be as motivating as explicit system-level policies, if not more so.⁸ In one district, for example, administrators informally held principals accountable for using data by frequently questioning them in meetings about how they planned to address declining test scores (Massell, 2001). Still others noted that intrinsic motivation plays a large part in spurring data use. Although Edison's Star Rating System may have motivated educators within Edison schools to pay attention to a wide range of data, the use of bonuses tied to ratings was not necessarily the driving force. Echoing other educators interviewed, one Edison principal explained, "I don't really think that, if a principal gets up every day, a bonus is what they're truly after. It's a nice ending to a year of hard work, but I don't think that's what really pushes them to reach that. I think it's the children" (Marsh et al., 2008, p. 443). In-depth research in one Edison school similarly found that teachers' intrinsic motivation, combined with external requirements to use benchmark assessments (along with capacity and supporting structures), promoted and sustained a culture of data use (Sutherland, 2004).

Finally, other studies indicate that there is a delicate, often difficult to achieve balance between "pushing" data users through accountability incentives and norms, and "pulling" them through professional development and other capacity-building efforts (terms used by Kronley & Handley, 2003, not reviewed herein). Several studies of interventions involving coaches or facilitators of data use demonstrate this tension (Datnow & Castellano, 2001; Gallimore et al., 2009; Goertz et al., 2009). For instance, Success for All facilitators were sometimes caught between monitoring the fidelity of model implementation (which includes the regular use of assessment data) and supporting teachers with the use of assessments and other elements of the reform. As such, facilitators had to work hard at developing the rapport and trust needed to serve as mentors and assistance providers to teachers who perceived them to be "quasi administrators" (Datnow & Castellano, 2001). As noted earlier, other studies indicate that administrators' use of interim assessment data can also tilt the balance away from support and more toward pressure in ways that inhibit educators' use of these data. For example, Philadelphia central office administrators required principals to complete data analysis protocols analyzing interim assessment results for their school and review them with their regional superintendents and other principals at monthly meetings. Although the process was intended to facilitate dialogue, researchers noted that "some educators viewed the public sharing of data as undermining the low-stakes, instructional focus of the interim assessments" (Goertz et al., 2009, p. 3).

Sustaining and Maintaining Sufficient Data Support Is a Significant Challenge

Researchers have found that maintaining not only the technical aspects of interventions but also the cultural supports (i.e., norms and expectations to use data) over time is very difficult (Ancess et al., 2007; Park & Datnow, 2009). Notably, several studies of BASRC found that leaders struggled to sustain support for data-driven inquiry when faced with considerable staff turnover in districts and schools, and changing expectations and contexts (Copland, 2003; Jaquith & McLaughlin, 2009; Porter & Snipes, 2006). Research on Edison Schools similarly found that staff attrition was a persistent challenge to sustaining capacity to implement the school design, including data use procedures and practices (Marsh et al., 2008).

Two studies, however, provide counterexamples of sustained data support. Copland's (2003) research on BASRC, for instance, identified some schools that continued to promote BASRC-related work despite experiencing leadership turnover by distributing the work among all staff and hiring personnel with a shared vision that included inquiry-based practice. Similarly, based on interviews with staff in 6 of 12 participating HRS secondary schools in Wales conducted 5 years after the formal project ended, researchers concluded that the "majority of schools continued using the high reliability principles" (p. 409) because the project built "capacity at the school level to continue educational development after the formal end of the project" (Stringfield et al., 2008, p. 424). Yet, they provide little substantiating evidence or details on what sustained use or capacity looked like.

Other studies have noted that maintaining depth of support—such as sustained, ongoing professional development for educators throughout a school or district during all stages of data use—is a related challenge. For example, several studies found that coaches are often unable to provide intensive support to all the teachers who need assistance in schools (Marsh, McCombs, & Martorell, 2010; Quint et al., 2008; Roehrig, Duggar, Moats, Glover, & Mincey, 2008). Similarly, studies of a PLC initiative concluded that inconsistency in the presence of facilitators to push teachers to critically analyze student work may have hindered the progress of these data teams and their ability to generate new thinking and practices (Nelson & Slavit, 2007; Slavit & Nelson, 2010).

Several studies have identified a lack of support for educators in the final stages of data use (leverage points 4 and 5). A recent national survey found that although most districts (90%) are providing schools with support on how to access student data from data systems, they are far less likely to provide all educators with training on how to use the system to

analyze student achievement or use the results to change their instructional practice (53%; Means et al., 2010). Other studies confirmed not only a dearth of, but also a great need for, additional support in this area. A study of data use in Philadelphia and one anonymous district concluded that “teachers need more professional development and support on interpreting data (e.g., diagnosing student error) and on connecting this evidence to specific instructional approaches and strategies” (Goertz et al., 2009, p. 241). The authors noted that adopting an aligned curriculum and tools is insufficient for enabling teachers to adjust instruction according to the data and that the most promising approaches are those that involve “regular, facilitated” support, such as facilitated teams of teachers who collectively examine data and instructional responses. Other studies suggest that coaches may be central to helping bridge this gap between data analysis and instruction (Marsh et al., 2010; Means et al., 2010; Roehrig et al., 2008). Researchers evaluating a portfolio assessment project similarly concluded that although the protocol and accompanying professional development influenced teachers’ thinking and practice in certain ways, more targeted and ongoing support, such as on-site coaching, was needed to fully translate the analysis of student work into fundamental changes in practice (Gearhart & Osmundson, 2009). A study of 23 reform-oriented districts also found that despite significant district efforts to improve knowledge and skills about data analysis and use, many educators struggled with “deriving data implications for their day-to-day instruction” and wanted help in making “feedback loops” more explicit (Massell & Goertz, 2002, p. 58). Case studies of 36 high-data-using schools from the national study cited earlier also indicate that educators wanted more training in how to interpret data and connect it to instructional practices—much more so than training on how to use data systems (Means et al., 2010).

A related theme that emerged in this review is that interventions vary widely in their depth of adoption and consistency of implementation (Copland, 2003; Jaquith & McLaughlin, 2009; Massell, 2001; Means et al., 2010; Nelson & Slavit, 2007; Stringfield et al., 2008). For example, researchers observed “developmental stages” of inquiry-based reform across participating BASRC schools, including those that were “novice,” “intermediate,” and “advanced” (Copland, 2003; Jaquith & McLaughlin, 2009). Similarly, Means et al. (2010) identified a “developmental progression” in case study schools’ data use activities. Other studies of system-level interventions found that transferring data-driven practice down to teachers did not occur consistently or with great depth (Jaquith & McLaughlin, 2009; Kerr et al., 2006; Moody & Dede, 2008; Porter & Snipes, 2006).

EVIDENCE ON INTERVENTION EFFECTS

Although some of the reviewed studies focused primarily on implementation, many commented on the effects of interventions on data users (including changes in attitudes, beliefs, knowledge, and behaviors), organizations (such as changes in norms and policies), and students (notably student achievement). As discussed further next, the quantity and quality of research supporting these findings vary widely.

EFFECTS ON DATA USERS

Researchers examine a wide range of measures to document effects of interventions on data users, including evidence of perceived usefulness; changes in attitudes, beliefs, and knowledge; frequency of data use following participation; and specific changes in practice. As discussed next, much of this research draws on self-reports from data users, and few studies rely on observations, authentic assessments of educator knowledge, or rigorous causal analyses.

Perceived Usefulness

Eight studies include evidence on the perceived usefulness of the intervention, which is generally favorable (Demie, 2003; Huffman & Kalnin, 2003; Kerr et al., 2006; Nelson & Slavit, 2007; Porter & Snipes, 2006; Stringfield et al., 2008; Supovitz & Weathers, 2004; Sutherland, 2004). For example, 90% of teachers who responded to a survey about a one-year collaborative inquiry seminar focused on math and science data agreed that it was a valuable experience (Huffman & Kalnin, 2003). Similarly, educators who participated in classroom- and school-level “walkthroughs” in two studies reported that they were useful learning experiences (Kerr et al., 2006; Supovitz & Weathers, 2004). In one of these studies, the perceived utility was significantly greater for educators who conducted the walkthroughs than for educators who were visited in the process: 53% of principals who were data collectors strongly agreed that the visit was useful to their faculty, compared with 26% of visited principals (Supovitz & Weathers, 2004). Also, U.K. administrators and teachers reported on customer satisfaction surveys that LEA tools and training were useful and easy to understand (Demie, 2003). As noted earlier, however, this study did not provide information about response rates or methods for administering surveys.

Attitudes, Beliefs, and Knowledge

Nine studies examined changes in data user attitudes, beliefs, or knowledge, and generally found mixed effects. The majority of this research, however, relied on self-reported data on effects, did not include pre-post measures and authentic assessments of knowledge, and fell far short of substantiating causal links between interventions and effects.

Positive Effects. Regarding attitudes, Massell (2001) noted that one of the “most striking effects” of state and district efforts to support data use was an increase in educators’ reports that *data were valuable* to practice (again, little detail or supporting evidence is provided). Similarly, survey results from three districts partnered with the IFL indicate that teachers were more likely to find data—test results and student work—useful for guiding instruction in their classrooms in the two districts that invested more in data support activities (Kerr et al., 2006).

As for beliefs, one study of Getting Results schools found that teachers involved in facilitated data inquiry within a team of colleagues began to attribute student achievement gains to their own teaching efforts—in contrast to the comparison teachers not involved in data teams, who attributed gains to external factors or student characteristics such as socioeconomic status (Gallimore et al., 2009). Other research on this same intervention found that participation altered teachers’ understanding of and expectations about the purposes of assessment data and “fostered an ‘improvement over time’ versus a ‘one-shot’ orientation for collecting, analyzing, and using data” (McDougall et al., 2007, p. 77).

As for knowledge and skills, more than 80% of teachers responding to a survey about a one-year math-science collaborative inquiry seminar agreed that the seminar increased their knowledge of teaching math and science, and more than 60% said it altered their philosophy of education (Huffman & Kalnin, 2003). Similarly, participants in a district-level strategic planning initiative reported learning how to ask different questions, how to think more holistically about all students in the district, and how to analyze data. One parent, for example, explained, “I look for different things now: Why did we score what we did [on the state test]? What’s the population of our school? I look more at the intricacies than I would ever have looked at before” (Marsh, 2007, p. 165).

Mixed or Negative Effects. Several studies surfaced negative effects of interventions on attitudes toward data, noting teacher frustration with the amount of time it takes to participate in these interventions and to use data more generally (Huffman & Kalnin, 2003; Kerr et al., 2006; Supovitz, 2006). In one study of five data inquiry teams, researchers observed variation in how teachers responded. Based on analyses of focus

groups with teachers, notes from facilitators, and audio and video recordings of meetings, they found that only two teams demonstrated changes in their understanding of inquiry-based teaching (Nelson & Slavit, 2007). Similarly, there was mixed evidence about the extent to which the Formative Assessment of Student Thinking in Reading (FAST-R) program in Boston—an intervention using coaches to support teachers' use of these reading assessment data—affected teachers' thinking. Although 70% of FAST-R teachers reported that the data helped them understand student thinking and answers, fewer than half said the data challenged their views of some students or that the data helped them question subgroup patterns in their classroom. And although the majority of teachers surveyed reported that the program contributed to their understanding of how data can be relevant to students' work (74%) and their understanding of different kinds of assessments and their uses (61%), teachers assigned to the control groups (which had other formative assessments and no support from a coach, but regular school or district professional development) were just as likely to report that professional development contributed to their understanding of how data could be used. Given low survey response rates, however, these findings should be interpreted cautiously.

Behavior—Frequency of Data Use and Changes in Practice

Although 20 studies reported on the effects of interventions on educator behavior or practice, once again, very few provide rigorous substantiating evidence or causal links. About half of this research found positive effects, whereas the other half suggested either mixed or negative behavioral responses.

Positive behavioral effects. Some studies have found positive effects on reported frequency of data use. For example, 92% of principals surveyed in one BASRC study reported that the use of data for decision making had moderately or substantially increased as a result of their school's participation with BASRC (Copland, 2003). These findings, however, must be judged cautiously because the author did not provide response rates, and another study found that reported use of assessment data by BASRC teachers remained relatively constant over time (Porter & Snipes, 2006). In several other studies, the reported frequency of use was correlated with level of support received. For example, Kerr et al. (2006) found that in two of the three study districts that invested more in data support activities, teachers and principals reported (on surveys) more extensive and frequent use of student achievement data and student work to guide instructional decisions. Teachers in these two districts were also much

more likely than counterparts in the third district to receive regular assistance with data analysis from their principals. Similarly, Anderson et al. (2010) found that principal reports on surveys about frequent use of data—including student achievement data, evaluation data, and research—were highly correlated with reports about district support for schools to use data for improvement planning. Although suggestive, these studies do not prove that the interventions increased data use.

Other studies provide limited, self-reported evidence of positive actions taken by data users in response to interventions (Denton et al., 2007; Marsh et al., 2010; Murnane et al., 2005; Quint et al., 2008; Supovitz, 2006; Sutherland, 2004).⁹ For example, more than 90% of participants in a one-year math-science collaborative inquiry seminar reported that it improved their classroom teaching and increased their collaboration with colleagues (Huffman & Kalnin, 2003). Limited evidence in one study indicates that participants in some data teams supported by university partners reported making school improvement decisions based on their analyses of data (e.g., new plans for professional development; Murnane et al., 2005). In another study, principals reported that data technology and tools influenced their practice and motivated teachers to learn new practices (Supovitz, 2006).

Finally, one study provides slightly more rigorous evidence of positive behavioral effects. Based on rubric-based coding of extensive observation data, interviews, and focus groups, evaluators found that teachers in four schools implementing the Getting Results (GR) model demonstrated qualitatively different procedures and behaviors in meetings than in the three comparison schools, including greater focus on student academics, broader analyses of data beyond test results, greater efforts to use and evaluate jointly developed instructional strategies, and more frequent accomplishment of instructional tasks (McDougall et al., 2007).

Mixed, negative, or no behavioral effects. In a study of the FAST-R program, more than 70% of teachers reported that participation contributed to their use of data to reflect on their instructional practice. Yet, researchers found no difference in the reported amount of time that FAST-R teachers spent looking at data compared with control group teachers. However, as noted earlier, the study relies heavily on a one-time survey measure of frequency with very low response rates (Quint et al., 2008).

Four other studies provide more rigorous evidence of effects on practice and again yield mixed results. Based on extensive interviews, observations, and surveys, researchers of an intervention involving formative assessments and training found that teachers used the data to alter *what* they taught but not *how* they taught. Teachers used the assessment results to decide what content to reteach and whom to target but did not make

fundamental changes in the way they taught students or content (Goertz et al., 2009). As noted earlier, these researchers concluded that more ongoing support is needed to facilitate these more substantive changes in practice.

Similarly, a study of a professional development effort using an assessment portfolio tool designed to help teams of teachers use and analyze formative assessments and student work relied on multiple data sources (portfolios, surveys, and focus groups) to assess effects on teacher practice (Gearhart & Osmundson, 2009). Although the researchers documented numerous positive changes that teachers made in their practice as a result of participation (e.g., greater focus on big ideas in setting goals, the use of new methods for interpreting student work for student understanding, and shifts away from simply using evidence to reteach to using evidence to target strategies for instructional improvement and provide meaningful feedback to teachers), they nonetheless identified gaps in teachers' ability to fully embrace the goals of the endeavor (e.g., understanding technical elements of assessment and learning curriculum-specific methods of using assessments to improve instruction).

In another study, middle and high school science and math teachers participating in one of five facilitated-inquiry teams demonstrated changes in the way they promoted critical thinking. After analyzing data from classroom observations and recordings of student questions in class, these teachers developed a classification scheme identifying higher order questions and new expectations for what they wanted to hear from students (Nelson & Slavit, 2007). In a related study of one inquiry team, teachers did not challenge each other to generalize to student understanding and referred to student thinking in vague terms (Slavit & Nelson, 2010).

Finally, seven studies demonstrate ways in which some data users respond to interventions in unintended ways, sometimes misusing data or not using them at all. For example, based on limited evidence, authors in one study found that for some data teams participating in their university-sponsored intervention, "the intention to 'game the system' seemed to take precedence over making constructive decisions in how students were taught" (Murnane et al., 2005, p. 279), including decisions to primarily improve test-taking skills rather than content knowledge and to focus on "bubble kids"—a practice that emerged in four other reviewed studies (Kerr et al., 2006; Marsh et al., 2008; Moody & Dede, 2008; Porter & Snipes, 2006). Other studies found that educators use data in ways that are "hasty" (Massell & Goertz, 2002) or "symbolic" (Coburn, 2010), or they make decisions without regard to data (Coburn, 2010).¹⁰

EFFECTS ON ORGANIZATIONS

Seven studies examined the effects of data support interventions on organizations—including effects on *organizational structure*, or formal rules, policies, and procedures, and on *organizational culture*, or the “set of taken-for-granted assumptions, shared beliefs, meanings, and values that form a kind of backdrop for action” (Smircich, 1985, p. 58). Taken as a whole, there is some evidence of interventions affecting culture and norms, but very few studies have examined or found evidence of structural changes occurring in organizations participating in these interventions.

Two studies reported changes in organizational structures and practices. Drawing on teacher self-reported survey data, one study found that 80% of seminar participants believed that the seminar helped their district use data, altered the way the district thinks about science and math, and improved curriculum and instruction in the district (Huffman & Kalnin, 2003). Relying on a wider array of data, including review of documents and interviews, other researchers found that schools involved in a whole-school adolescent health intervention involving data teams were much more likely to introduce or make changes to written policies (e.g., to prevent bullying, promote staff–student communication) and strategies (e.g., peer mediation) to improve school and student outcomes as compared with comparison schools (Bond et al., 2001).

In six studies, researchers found that interventions affected organizational culture and norms (Corcoran & Lawrence, 2003; Huffman & Kalnin, 2003; Kerr et al., 2006; Jaquith & McLaughlin, 2009; Supovitz, 2006; Sutherland, 2004). Again, in some cases, these studies provide either weak or little credible evidence substantiating these effects. For example, districts partnered for more than a decade with the Merck Institute for Science Education Partnership demonstrated changes in norms of inquiry systemwide (Corcoran & Lawrence, 2003). Based on summary analyses of longitudinal survey, observation, and interview data, researchers reported—with few supporting details—that in partner districts, practice became more public; teachers were more willing to discuss their instruction, share ideas and materials, and examine performance data; and the incidence of peer observations, collective review of student work, and coteaching increased. Similarly, investments in districtwide systems supporting data use in Duvall County were reported to facilitate organizational learning, a culture of inquiry, and a collective sense of accountability. For example, the researcher noted that the district’s use of its “snapshot” data system helped define and communicate expectations to educators systemwide about high-quality practice and

encouraged educators at all levels to collectively examine, discuss, and receive feedback on their practice—norms and practices said to be less prevalent prior to the implementation of this reform (Supovitz, 2006).

EFFECTS ON STUDENTS

Six studies reviewed directly examined the effects of data support on student outcomes.¹¹ Four studies referred to the overall effects of a reform initiative without isolating the data-related supports per se. Two of them found positive effects. McDougall et al. (2007) reported that the nine elementary schools using the Getting Results model showed significantly greater gains in test score averages (across Grades 2–5 and reading, math, language, and spelling subtests of the Stanford 9) than six comparison schools across 5 years. Researchers reported an effect size of 0.75 but provided few details and referred readers to an evaluation report for more information. Similarly, research on 12 Welsh secondary schools implementing the HRS model found positive student outcomes after 4 years in the program and 5 years postintervention. More specifically, the rate of gain in the number of students passing five or more national exams was much greater for HRS schools than the national average at the 4- and 9-year marks (Stringfield et al., 2008). In contrast, a study of BASRC’s “focal strategy” of district-level coaching and broader supports for inquiry-based practice found that improvements in reading achievement in participating districts were not substantially different than in similar local districts (Porter & Snipes, 2006). Similarly, the randomized controlled trial of Boston’s FAST-R program found no significant effect of the assessments and coaching on reading test scores, but researchers noted that the coaching intervention was very low in intensity (Quint et al., 2008).

Some of these studies, and others, attempt to relate implementation intensity with student outcomes. For example, BASRC evaluators found that schools with higher levels of BASRC-related practices did not have higher levels of achievement (Porter & Snipes, 2006). In contrast, researcher found that HRS schools using more of the model’s practices, including the use of various types of data for decisions and monitoring, showed significantly greater achievement gains than those using less (Reynolds, Stringfield, & Schaffer, 2006). Similarly, one study of middle school reading coaches found that more frequent time spent analyzing student data with teachers was associated with higher student achievement (Marsh et al., 2010). None of these studies, however, provides strong causal evidence linking data support and student effects.

In sum, among the limited number of studies examining the effect of

interventions on student achievement, there is mixed evidence of effects. Similarly, there is limited research on student achievement effects of data use more generally, and in fact, little supporting evidence of positive effects. A 2009 panel of experts and a review of research found a low level of evidence supporting any causal claims between the use of data and student outcomes—due in part to a lack of rigorous experimental studies, the difficulty of isolating the contributions of specific elements of data use practices, and the possibility that it is too early in the practice of these reforms to have accumulated a valid body of evidence (Hamilton et al., 2009).

CONDITIONS INFLUENCING DATA USE INTERVENTIONS

Even the best designed interventions may not succeed if the contextual conditions are not conducive to implementation. Many studies reviewed suggest a common set of conditions that appear to affect the implementation and effects of data use interventions, and many of these overlap with conditions found to facilitate or constrain data use more generally. Given the variation in research quality, many of these findings are best understood as suggestive. I start with conditions related to the intervention (capacity of intervener, data properties), move to broader contextual conditions (leadership, organizational structure, time, policy context), and end with conditions related to individuals (interpersonal relationships and belief and knowledge).

CHARACTERISTICS OF THE INTERVENTION

Capacity of Intervener

Twelve studies indicate that the capacity of the organizers of interventions and their staff can greatly influence the outcome of these efforts. Many studies point to the expertise of interveners, noting that staff delivering services must have strong data literacy skills (e.g., knowledge of how to examine multiple measures, synthesize data, and draw inferences; Demie, 2003), deep knowledge of teaching and learning (Park & Datnow, 2009), and an ability to work with adults (Marsh et al., 2010). Many studies also have found that support providers often vary in their level of skills and knowledge, which affects interactions with educators (Denton et al., 2007; Honig & Ikemoto, 2008; Marsh et al., 2010). Nevertheless, none of these studies directly and objectively measures the knowledge and skills of interveners but instead rely on reports from interveners and educators being supported.

Other studies suggest that possessing internal and external expertise matters. In this research, internal expertise refers to local knowledge and skills relevant to the specific context in which educators practice, and external expertise refers to a broader set of knowledge and skills usually brought in by individuals who are not as familiar with or employed by the particular school or district. One study found that schools with lower levels of data use tended to rely on external expertise to guide and support data use, whereas schools with higher levels of data use exhibited widely distributed internal expertise (Anderson et al., 2010). Other researchers identified a tension among intermediaries being perceived by educators as local versus “system” level experts (Honig, 2004). A study of Getting Results schools found that external support providers initially perceived to be “outsiders” over time “morphed into a more internal form of assistance as educators began to interact with GR staff as ‘one of our own’” (McDougall et al., 2007, p. 86).

Staff attrition and turnover has also been found to create problems for sustaining data support over time (Copland, 2003; Marsh et al., 2008; Porter & Snipes, 2008). Studies also found capacity limitations of intermediary organizations trying to support data use on a large scale. For example, studies of BASRC indicate that leaders had insufficient capacity to support change at that scope and level of complexity (Jaquith & McLaughlin, 2009) and that the reform effort suffered from limits in the ability of BASRC to simultaneously support educators at the district and school levels (Porter & Snipes, 2006).

Properties of Data

Several studies have asserted that characteristics of the focal data relate to intervention success or failure. Some researchers asserted that comparative data facilitate deep conversations and learning, but again, they often failed to provide rigorous evidence supporting this connection (Huffman & Kalnin, 2003; Massell, 2001; Supovitz & Weathers, 2004). For example, Huffman and Kalnin (2003) asserted, without a lot of supporting evidence, that the “intentional juxtaposition” of local and international math and science test results in the professional development program “contributed to the participants developing a sense of ownership” and “encouraged substantive conversations among team members about the interactions between structural and instructional factors that may affect mathematics and science achievement” (p. 579). Others similarly pointed to the value of including high-quality data (Corcoran & Lawrence, 2003) and multiple types of data in the process (Massell, 2001; Park & Datnow, 2009) and based these findings on reports from partici-

pants that these data qualities led to a more comprehensive and valid understanding of problems and potential solutions. In contrast, one study noted that multiple measures sometimes yield competing views of achievement, which complicates teacher efforts to derive data implications and next steps (Massell & Goertz, 2002).

BROADER CONTEXT

Leadership

Studies of coaching (Marsh et al., 2010; Roehrig et al., 2008), intermediary and school reform organizations (Copland, 2003; Gallimore et al., 2009; McDougall et al., 2007; Stringfield et al., 2008), and school systems (Anderson et al., 2010; Goertz et al., 2009; Marsh et al., 2008; Massell, 2001; Park & Datnow, 2009; Sutherland, 2004) have found that principals and central office administrators provide much-needed vision, modeling, and direction, enabling data-use interventions to flourish. Nevertheless, as noted throughout this article, the quality of evidence supporting the link between leadership and support for data use varies greatly across the studies reviewed. For example, although Stringfield et al. (2008) reported that the “the role of the head [principal of HRS schools in Wales] was pivotal in schools’ progress” in achieving academic gains, it provides little supporting evidence in this particular document (p. 425). Similarly, a district intermediary study provides suggestive evidence of this relationship, finding a correlation between districts in which teachers reported more frequent use of data to adapt teaching practices and those in which teachers reported receiving regular support from their principal with these endeavors (Kerr et al., 2006). Drawing on a slightly richer evidence base—survey data from a limited sample of teachers and principals, along with interview and observation data analyzed through the lens of a conceptual framework—Copland (2003) described the ways in which principals and teacher leaders facilitated the work of the intermediary organization BASRC and promoted involvement in inquiry and data use.

Four studies identified distributed leadership as a key facilitator of data use (Copland, 2003; Gallimore et al., 2009; McDougall et al., 2007; Sutherland, 2004). For example, Sutherland (2004) reported that the multiple structured leadership roles (e.g., curriculum coordinators, house leads) promoted by the Edison Schools model maintained a “unified focus on data use” and nurtured a culture of data use for improvement (p. 289). Gallimore et al. (2009) similarly found that using teacher-facilitators for inquiry teams freed up coaches and content

experts to serve as knowledge “resources” and principals “to circulate and provide appropriate support and accountability for multiple teams and facilitators” (p. 548). Finally, researchers examining HRSs asserted that well-managed leadership succession—filling open positions with individuals from other high-implementing schools—played a critical role in ensuring sustained implementation of reforms and student achievement gains (Stringfield et al., 2008).

Organizational Structure

A handful of other studies found that certain organizational structures complicated the use of data. Two identified the limitations of year-round schedules. In a study of BASRC, a district’s adoption of a year-round calendar created irregular breaks that complicated the tracking of student data and meetings with staff to discuss progress (Copland, 2003). Evaluators of Getting Results schools also found that year-round, multi-track schedules inhibited planning and communication, detracted from the focus on data in meetings, and compromised teacher buy-in because one third of staff were always absent when decisions were made (McDougall et al., 2007). According to two district studies, divisions and “silos” within the central office fractured reform efforts and limited opportunities to develop shared understandings around data (Coburn, 2010; Moody & Dede, 2008).

Time

Six studies identified lack of time as a barrier to data use interventions (Coburn, 2010; Gallimore et al., 2009; Marsh et al., 2010; Means et al., 2010; Quint et al., 2008; Slavit & Nelson, 2010). For example, 92% of districts in a nationally representative survey reported that lack of time was a barrier to expanding the use of data-driven decision making in their district (Means et al., 2010). Similarly, a study of one inquiry team found that 30-minute meeting times inhibited teachers’ ability to have prolonged discussions and contributed to their inability to make progress using data to guide instruction (Slavit & Nelson, 2010).

Other literature provides positive examples of interventions that ensured adequate time was available for this process (Goertz et al., 2009; McDougall et al., 2007; Means et al., 2010; Park & Datnow, 2009). For example, 30 of the 36 high-data-use case study schools studied in Means et al. (2010) regularly set aside time for examination of data (e.g., in weekly team or department meetings). Leaders in Philadelphia built in a

week of time after the administration of its interim assessments to allow for review and/or extended development of topics identified by the analysis of results as needing attention (Goertz et al., 2009). Similarly, Duvall leaders built in time for collective discussion and feedback of snapshot data to principals (Supovitz, 2006).

Policy Context

According to five studies, misalignment between data interventions and other policies weakens intervention effects. For example, misalignment between interim assessments and curriculum content and pacing was a significant obstacle to educators' ability to use assessment results to guide practice in four studies. Researchers found that without the flexibility or authority to adjust instruction or curriculum according to what is discovered after analyzing the data, educators could not effectively act on the data (Kerr et al., 2006; Means et al., 2010; Porter & Snipes, 2006; Wohlstetter et al., 2008). Similarly, teachers in one inquiry team complained that despite the enhanced skills and sense of professionalism gained from participating in workshops, they did not "have the authority to get the work finished" and enact policy changes once the formal program ended (Huffman & Kalninm, 2003, p. 577).

INDIVIDUAL RELATIONSHIPS AND CHARACTERISTICS

Interpersonal Trust and Relationships

Six studies highlighted the importance of data user relationships and trust, particularly in interventions involving collective examination of data (Copland, 2003; Datnow & Castellano, 2001; Marsh, 2007; Means et al., 2010; Nelson & Slavit, 2007; Park & Datnow, 2009). These studies suggested that groups without trusting relationships experience less progress in using data and attaining desired outcomes than those with trusting relationships. For example, Nelson and Slavit (2007) found that two of five teacher teams demonstrating greater progress in implementing an inquiry process had well-established professional relationships prior to the program, which teachers felt were "conducive to opening their classroom practices up to group examination" (p. 32). Similarly, researchers examining 36 high-data-use schools concluded, "A persistent theme across schools in which teachers look at data collaboratively is the importance of their sense that their colleagues respect them and their trust that the data will not be used to punish them" (Means et al, 2010, p. 71).

Data User Beliefs, Skills, and Knowledge

Finally, a few studies found that individual-level characteristics influence the process of using and supporting data. Three note that individuals' preexisting beliefs and priorities shape how they respond to data-support interventions and explain why implementation may not align with intended goals or design (Coburn, 2010; Honig & Ikemoto, 2008; Kerr et al., 2006). For example, district administrators working with external university partners often interpreted data and research through the lens of beliefs (e.g., views about effective instruction), frequently arriving at different implications or ignoring the data altogether (Coburn, 2010). Another study identified the critical importance of educators' perceptions about the validity of the data (Kerr et al., 2006).¹²

Three studies pointed to individual skills and knowledge as critical mediators of data use. For example, more than half of districts in a nationally representative survey reported that lack of teacher preparation on how to use data (56%) and lack of technical skills of staff to use data systems (69%) was a major barrier to expanding the use of data-driven decision making in their district (Means et al., 2010). Another study found a correlation between districts in which teachers reported more frequent use of data to adapt teaching practices and those in which teachers reported feeling prepared to interpret and use test data (Kerr et al., 2006). Massell (2001) also noted that one barrier to district efforts to support data use is the fact that teachers are not taught data analysis skills in their preparation programs.

Ultimately, this collective body of research provides a broad understanding that these intervention, contextual, and individual conditions matter but leaves many questions unanswered about to what degree they matter to implementation and outcomes and in what specific ways. I return to some of these questions in the next section.

SUMMARY AND DIRECTIONS FOR FUTURE RESEARCH

This review of literature suggests that supporting educators' use of data is a complex endeavor fraught with multiple challenges as well as opportunities. As noted throughout, this research base is limited in quantity and quality. Many of the studies reviewed provide weak or little credible evidence substantiating findings and rely heavily on self-reported data rather than rigorous causal analyses to document intervention effects. These limitations are important to keep in mind when considering the overall insights and the multitude of unanswered questions.

Regardless of whether an intervention is designed as part of a comprehensive reform initiative or as a stand-alone professional development effort, some evidence suggests that the process thrives when interventions ensure that data are easy to understand and usable; include norms and structures promoting the safety and confidentiality of data and data discussions; target multiple leverage points; and involve opportunities for cross-site (or level) collaboration. The research further indicates that interventions frequently face several persistent challenges, such as developing support that is generic but also customized, and conceptual but also concrete, establishing the appropriate role of accountability, and sustaining and maintaining depth of data support over time.

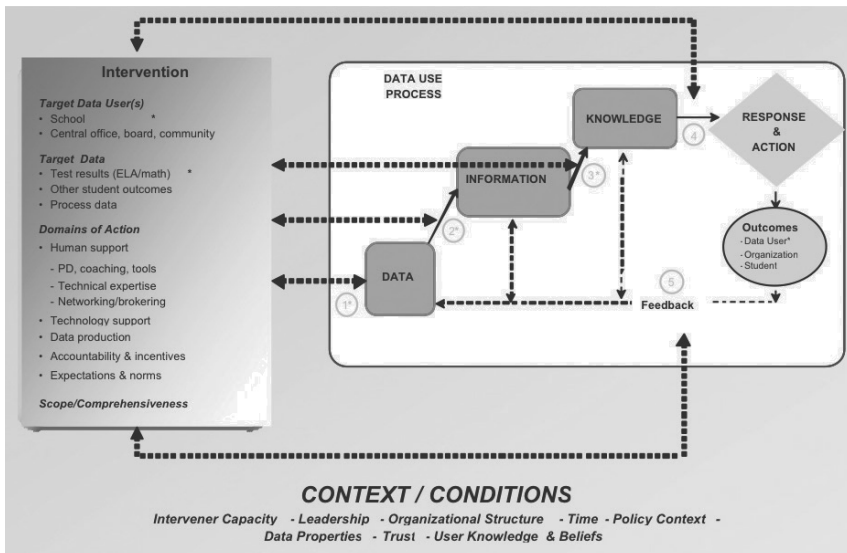
Ultimately, there are mixed findings and levels of research evidence on the effects of data support interventions. Relatively more is known about the effects of interventions on educators' knowledge, skills, and practice than on the effects on organizations or on the long-term effects on student achievement. Nevertheless, much of the research on intermediate effects on data users relies on self-reported data and lacks direct observational measures or causal evidence of changes in educators' knowledge or practice resulting from participation in interventions. And although findings from a handful of studies provide mixed evidence of intervention effects on student achievement, to date, intensive, well-designed data support interventions have not been studied carefully or with much attention to measuring longer term outcomes. Further, very little is known about the effects of certain types of interventions relative to others.

Finally, this research begins to suggest a set of common conditions that influence the implementation and effects of data support interventions, including intervention characteristics (capacity, data properties), broader context (leadership, organizational structure), and individual relationships and characteristics (trust, beliefs, and knowledge). Again, the amount and quality of evidence supporting these findings vary considerably, and much remains unanswered regarding the contextual conditions needed to facilitate data support.

In sum, the terrain of the research reviewed and an overall framework for understanding data support interventions can be illustrated in Figure 2. The Interventions box at the left highlights several dimensions that current research indicates are important for understanding the process of supporting educators' use of data, including the target data user and types of data, domains of action, and scope or comprehensiveness. The bold dotted arrows illustrate the multiple leverage points at which interventions may support data use, such as accessing data, organizing them into information, combining information with expertise and understandings

in ways that build knowledge, helping users respond to and act on new knowledge, and providing feedback on the process. Finally, the outer box suggests contextual conditions that may mediate the process of data support. Although the collective body of literature reviewed touches on all these domains and dimensions, the quality and quantity of research supporting each vary greatly. Words that have an asterisk (*) indicate areas with greater amounts of credible, supporting evidence.

Figure 2. Mapping the terrain of research on data support interventions



In the next section, I elaborate on areas within this framework that are worthy of more research. I developed this research agenda based on three guiding questions: (1) Which areas are understudied or currently lacking in rigorous evidence (priority on those with the least amount of high-quality research)? (2) Where is there demand for further knowledge from practitioners and policy makers? (3) How can future studies address methodological weaknesses of the current research base?

FUTURE RESEARCH

First, not enough attention has been paid to the organizational- and student-level outcomes of data support interventions. Given the significant national push for data-driven practice and positive claims asserted by data advocates, more research is needed on interventions intending to

support this work to determine whether they help achieve intended outcomes. For example, one might use experimental designs to randomly assign data users to a variety of treatments, including no data, data without intervention, and data with intervention, and measure the effects at all levels, including school, teacher, and student. Future research on intervention effects on participating organizations should be longitudinal and examine the evolution of structures, policies, and culture during and after interventions. And although studies provide some evidence of effects on data users (relying primarily on self-reported data), a lot more could be learned through direct observation, pre-post measures of knowledge and skills, and comparison groups about how data users respond in their day-to-day work to participation in these activities. Once again, longitudinal studies are also needed to examine not only the effects on data users over time but also to identify means of sustaining data use practices once interventions have ended.

Second, there is a dearth of comparative analysis of interventions. Although a few authors compared the implementation of certain reform efforts across districts or schools, even fewer compared different types of interventions. Such comparisons could provide systematic information about what occurs in these interventions and how data users interact with interveners and data. They could also identify strategies and features of interventions resulting in better or worse outcomes for data users and students, and conditions that matter most. For example, one could compare different types of human supports, such as a data coach, literacy coach, and inquiry group, and ask: How does the support from a data coach with technical expertise compare with that from a literacy coach with content-area expertise or from a data team that brings distributed expertise? What type of knowledge and expertise among these facilitators/support providers helps with data analysis and translating new knowledge into practice? What particular practices or strategies are more effective and under what conditions? One could also randomly assign or find sites that vary the leverage points around which interventions are provided to examine potential differences in the processes, strategies, and outcomes of supporting different stages of the process (e.g., organizing or interpreting data versus translating knowledge to action). Such studies could answer the questions: Do certain interventions offer greater “purchase” at different leverage points? For example, does technology have more of an impact on organizing data versus taking action? Do human supports result in more positive data user behaviors than accountability measures? One could also seek out variation in types of focal data (e.g., student work versus test scores; mathematics versus ELA test scores; open-ended versus multiple-choice test results), types of users

(e.g., teachers versus administrators), and settings (e.g., elementary versus secondary; smaller versus larger organizations).

Given the widespread finding surfaced in this review that data users often lack capacity to use data and want or lack support to help move from knowledge to action, it would also behoove the research community to pay particular attention to interventions targeting leverage point 4. Researchers could examine the training provided around data use in teacher and administrator education and preparation programs, as well as in-service programs designed to help educators bridge data and practice. In light of the current economic climate and limited resources facing policy makers, future research should also consider the costs and benefits of these interventions. As the field gains more evidence concerning the effects of interventions on teachers, schools, and students, researchers can work to determine whether each intervention's benefits are worth the cost when compared with other interventions or investments.

As for conditions enabling data support, greater attention to the quality of data and leadership practices could also enhance this research base. Aside from some attention to users' perceptions of the validity of data, few studies have carefully examined the ways in which properties of the data (e.g., reliability, validity) affect users' responses to data and the processes and outcomes of interventions. More research is also needed to identify specific leadership practices and actions that matter most for outcomes desired. And although we generally know that other conditions matter, more could be learned about time (What is the minimum amount of time needed to ensure data use and how should it be structured?), interpersonal trust (What are strategies for developing trust over time?), and intervener and data user capacity (How important are certain skills and knowledge to actual use of data, and what specific skill sets are required?).

The scope of research on this topic might also be expanded to include a broader range of data types and target data users. Although three fourths of the studies reviewed focused on mathematics and ELA test results, significantly fewer examined data on learning in other subject areas, other student outcomes (e.g., student engagement), parent satisfaction, and processes, such as the quality of instruction or program implementation. Elementary schools are also disproportionately represented in this literature, and more could be learned about how to support high school educators, central office administrators, school board members, and students in using data to guide their work.

Finally, future research could address some of the methodological weaknesses identified throughout this review. First, with the exception of

a few studies, much of the existing research is atheoretical, lacking frameworks for understanding the mechanisms by which interventions produce documented results. Some promising frameworks include organizational and sociocultural learning theory (see, for example, Honig & Ikemoto, 2008) as well as theoretical lenses that take into consideration political dimensions of data use (for a review of this literature, see Henig, 2012, this issue). For example, sociocultural learning theory literature might better define the domains of support or activities that one might expect to observe of a coach working with teachers to support data use—drawing attention to the importance of modeling and making thinking “visible,” for instance. Second, future studies might also strive for greater generalizability, examining interventions in settings in which participants have not volunteered to participate, which may come closer to generating knowledge about how to support a typical teacher or administrator with data use. Third, researchers should seek to provide more credible evidence at all levels, using larger and more representative samples, more authentic measures that move beyond self-reports, triangulation of methods and sources, and clear documentation of data collection and analysis methods. Finally, researchers should explore research designs that allow for more rigorous causal inferences about intervention effects and conditions affecting outcomes.

Notes

1. Many of these single, system, or intermediary case studies include multiple embedded school case studies.

2. Unlike the professional development initiatives that generally focus exclusively on data support, many of the system or intermediary initiatives are not solely focused on supporting data use but instead include expectations and supports for data use as one component of a broader set of improvement strategies (e.g., a school reform model not only requires the use of assessments and provides related training but also contains specific structural arrangements, curricula, and so on). Whenever possible, I have tried to isolate study findings about the data support elements of these broader reform models but acknowledge that in some cases, study findings may apply more broadly to the overall set of strategies embodied by the reforms.

3. A few articles drew on larger studies that employed quasi-experimental or experimental designs, but are not included in these counts because the article reviewed did not draw on these analyses (e.g., Denton et al., 2007).

4. It is possible that some of these studies employed stronger methods than those described in the documents reviewed. For example, in some cases, authors were reflecting on years of research that may have been documented in past reports that included more explicit details of methods and more evidence than what was included in the reviewed document (e.g., Corcoran & Lawrence, 2003; Jaquith & McLaughlin, 2009). My assessments are based solely on what was presented in the documents reviewed.

5. Given the predominant use of qualitative methods in the reviewed documents, I draw on the concepts of “trustworthiness” and “credibility” to describe the quality of

research and rely on commonly cited criteria of clear exposition of data collection and analysis methods, triangulation, thick descriptions/authentic evidence, and confirmability (Lincoln & Guba, 1985; Patton, 2002). I have included several studies that I authored or coauthored and recognize the challenge of judging the quality of one's own work but have done my best to apply the same criteria and level of scrutiny to my work as to other pieces reviewed.

6. A few authors developed a framework as an outcome of their research (e.g., Roehrig, Duggar, Moats, Glover, & Mincey, 2008; Supovitz, 2006).

7. Other studies, not reviewed herein, have described a similar set of tensions among the work of reform support organizations (Bodilly, 1998) and theory-based intermediaries more generally (McLaughlin & Mitra, 2001).

8. O'Day (2002) made a similar argument about the limits of outcome-based, bureaucratic accountability policy more generally and found that a combination of bureaucratic and professional accountability is more likely to focus attention on information relevant to teaching and learning and to motivate individuals to use and respond to data.

9. Past research has called into question the accuracy of teachers' self-reports of behavior and found little correlation between observational data and self-reported measures of instruction (e.g., Hook & Rosenshine, 1979). And although some have demonstrated that teachers have the ability to accurately report on their instruction, the question of social desirability bias remains (Newfield, 1980).

10. Other studies examining the use of research similarly find that individuals tend to ignore or symbolically use this type of information (e.g., Feldman & March, 1981; Weiss, 1980).

11. Some authors refer to other studies in which the broader reform effort or experiment examined student outcomes (e.g., Bond et al., 2001, refers to unpublished findings on an earlier randomized controlled trial; Marsh et al., 2008, refers to Gill et al., 2005, for broader achievement analysis of Edison schools). These studies, however, do not isolate the effects of data support interventions.

12. Other research not reviewed shows that perceptions about data validity can vary greatly (Coburn & Talbert, 2006) and that form of data (e.g., multiple-choice, open-ended) may also affect its use and perceptions of validity and reliability (see for example, Koretz, Mitchell, Barron & Keith, 1996). Other studies similarly highlight the role of interpretative processes and cognition in shaping policy implementation (e.g., Spillane, 2000).

References

- Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, 16, 3–9.
- Ancess, J., Barnett, E., & Allen, D. (2007) Using research to inform the practice of teachers, schools, and school reform organizations. *Theory Into Practice*, 46(4), 325–333.
- Anderson, S. E., Leithwood, K., & Strauss, T. (2010). Leading data use in schools: Organizational conditions and practices at the school and district levels. *Leadership and Policy in Schools*, 9(3), 292–327.
- Bodilly, S. J. (1998). *Lessons from the New American Schools' scale-up phase: Prospects for bringing design to multiple schools*. Santa Monica, CA: RAND.
- Bond, L., Glover, S., Godfrey, C., Butler, H., & Patton, G. C. (2001). Building capacity for system-level change in schools: Lessons from the Gatehouse Project. *Health Education Behavior*, 28(3), 368–383.
- Choppin, J. (2002, April). *Data use in practice: Examples from the school level*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

- Coburn, C. E. (2010). Partnership for district reform: The challenges of evidence use in a major urban district. In C. E. Coburn & M. K. Stein (Eds.), *Research and practice in education: Building alliances, bridging the divide* (pp. 167–182). New York, NY: Rowman & Littlefield.
- Coburn, C. E., Honig, M. I., & Stein, M. K. (2009). What's the evidence on district's use of evidence? In J. Bransford, D. J. Stipek, N. J. Vye, L. Gomez, & D. Lam (Eds.), *Educational improvement: What makes it happen and why?* (pp. 67–86). Cambridge, MA: Harvard Educational Press.
- Coburn, C. E., & Talbert, J. E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education* 112(4), 469–495.
- Copland, M. (2003). Leadership of inquiry: Building and sustaining capacity for school improvement. *Educational Evaluation and Policy Analysis*, 25, 375–395.
- Corcoran, T., & Lawrence, N. (2003). *Changing district culture and capacity: The impact of the Merck Institute for Science Education Partnership* (CPRE Research Report Series RR-054). Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.
- Data Quality Campaign. (2009). *The next step: Using longitudinal data systems to improve student success*. Retrieved from <http://www.dataqualitycampaign.org/files/NextStep.pdf>
- Datnow, A., & Castellano, M. (2001). Managing and guiding school reform: Leadership in Success for All schools. *Educational Administration Quarterly*, 37(2), 219–249.
- Demie, F. (2003). Using value-added data for school self-evaluation: A case study of practice in inner-city schools. *School Leadership & Management*, 23(4), 445–467.
- Denton, C. A., Swanson, E. A., & Mathes, P. G. (2007). Assessment-based instructional coaching provided to reading intervention teachers. *Reading and Writing: An Interdisciplinary Journal*, 20(6), 569–590.
- Feldman, J., & Tung, R. (2001, April). *Whole school reform: How schools use the data-based inquiry and decision making process*. Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
- Feldman, M., & March, J. (1981). Information in organizations as signal and symbol. *Administrative Science Quarterly*, 26, 171–186.
- Gallimore, R., Ermeling, B.A., Saunders, B., & Goldenberg, C. (2009). Moving the learning of teaching closer to practice: Teacher education implications of school-based inquiry teams. *Elementary School Journal*, 109(5), 537–553.
- Gearhart, M., & Osmundson, E. (2009). Assessment portfolios as opportunities for teacher learning. *Educational Assessment*, 14(1), 1–24.
- Gill, B., Hamilton, L., Lockwood, J. R., Marsh, J., Zimmer, R., Hill, D., & Pribesh, S. (2005). *Inspiration, perspiration, and time: Operations and achievement in Edison Schools* (No. MG-351-EDU). Santa Monica, CA: RAND. Retrieved from http://www.rand.org/pubs/monographs/2005/RAND_MG351.pdf
- Goertz, M., Olah, L., & Riggan, M. (2009). *From testing to teaching: The use of interim assessments in classroom instruction* (CPRE Research Report No. RR-65). Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.
- Halverson, R., & Shapiro, B. (in press). *Technologies for learning and learners: How data are (and are not) changing schools* (WCER Working Paper). Madison: Wisconsin Center for Education Research.
- Hamilton, L., Halverson, R., Jackson, S., Mandinach, E., Supovitz, J., & Wayman, J. (2009). *Using student achievement data to support instructional decision making* (NCEE 2009-4067). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Henig, J. R. (2012). The politics of data use. *Teachers College Record*, 114(11)

- Honig, M. I. (2004). The new middle management: Intermediary organizations in education policy implementation, *Educational Evaluation and Policy Analysis*, 26(1), 65–87.
- Honig, M. I., & Ikemoto, G. (2008). Adaptive assistance for learning improvement efforts: The case of the Institute for Learning. *Peabody Journal of Education*, 83(3), 328–363.
- Hook, C. M., & Rosenshine, B. V. (1979). Accuracy of teacher reports of their classroom behavior. *Review of Educational Research*, 49, 1–11.
- Huffman, D., & Kalnin, J. (2003). Collaborative inquiry to make data-based decisions in schools. *Teaching and Teacher Education*, 19(6), 569–580.
- Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record*, 106, 1258–1287.
- Jaquith, A., & McLaughlin, M. (2009). A temporary, intermediary organization at the helm of regional education reform: Lessons from the Bay Area School Reform Collaborative. *Second International Handbook of Educational Change, Springer International Handbooks of Education*, 23(1), 85–103.
- Kerr, K. A., Marsh, J. A., Ikemoto, G. S., Darilek, H., & Barney, H. (2006). Districtwide strategies to promote data use for instructional improvement. *American Journal of Education*, 112, 496–520.
- Koretz, D., Mitchell, K., Barron, S., & Keith, S. (1996). *Perceived effects of the Maryland School Performance Assessment Program*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing and RAND.
- Kronley, R. A., & Handley, C. (2003). Reforming relationships: School districts, external organizations, and systemic change. Report prepared for School Communities That Work: A National Task Force on the Future of Urban Districts. Providence, RI: Annenberg Institute for School Reform at Brown University.
- Lachat, M. A., & Smith, S. (2005). Practices that support data use in urban high schools. *Journal of Education for Students Placed at Risk*, 10(3), 333–349.
- Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage.
- Love, N. (2004). Taking data to new depths. *Journal of Staff Development*, 25(4), 22–26.
- Mandinach, E., Honey, M., Light, D., & Brunner, C. (2008). *A conceptual framework for data-driven decision-making*. In E. B. Mandinach & M. Honey (Eds.), *Data-driven school improvement: Linking data and learning* (pp. 13–31). New York, NY: Teachers College Press.
- Marsh, J. (2007). *Democratic dilemmas: Joint work, education politics, and community*. Albany, NY: SUNY Press.
- Marsh, J., Hamilton, L., & Gill, B. (2008). Assistance and accountability in externally managed schools: The case of Edison Schools, Inc. *Peabody Journal of Education*, 83(3), 423–458.
- Marsh, J., McCombs, J. S., & Martorell, F. (2010). How instructional coaches support data-driven decision making: Policy implementation and effects in Florida middle schools. *Educational Policy*, 24(6), 872–907.
- Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). *Making sense of data-driven decision making in education: Evidence from recent RAND research* (No. OP-170-EDU). Santa Monica, CA: RAND Corporation.
- Mason, S. (2002). *Turning data into knowledge: Lessons from six Milwaukee Public Schools*. Madison: Wisconsin Center for Education Research.
- Massell, D. (2001). The theory and practice of using data to build capacity: State and local strategies and their effects. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 148–169). Chicago, IL: University of Chicago Press.
- Massell, D., & Goertz, M. E. (2002). District strategies for building instructional capacity. In A. M. Hightower, M. S. Knapp, J. A. Marsh & M. W. McLaughlin (Eds.), *School districts and instructional renewal* (pp. 43–60). New York, NY: Teachers College Press.

- McDougall, D., Saunders, W., & Goldenberg, C. (2007). Inside the black box of school reform: Explaining the how and why of change at Getting Results schools. *International Journal of Disability, Development and Education*, 54(1), 51–89.
- McLaughlin, M. W., & Mitra, D. (2001). Theory-based change and change-based theory: Going deeper, going broader. *Journal of Educational Change*, 2, 301–323.
- Means, B., Padilla, C., & Gallagher, L. (2010) *Use of education data at the local level: From accountability to instructional improvement*. Washington, DC: U.S. Department of Education, Office of Planning, Evaluation, and Policy Development
- Moody, L., & Dede, C. (2008). Models of data-based decision making: A case study of the Milwaukee Public Schools. In E. Mandinach & M. Honey (Eds.), *Linking data and learning* (pp. 233–254). New York, NY: Teachers College Press.
- Murnane, R., Sharkey, N., & Boudett, K. (2005) Using student-assessment results to improve instruction: Lessons from a workshop. *Journal of Education for Students Placed at Risk*, 10(3), 269–280.
- Nelson, T. H., & Slavit, D. (2007). Collaborative inquiry among science and mathematics teachers in the USA: Professional learning experiences through cross-grade, cross-discipline dialogue. *Professional Development in Education*, 33(1), 23–39.
- Newfield, J. (1980). Accuracy of teacher reports: Reports and observations of specific classroom behaviors. *Journal of Educational Research*, 74(2), 78–82.
- O’Day, J. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review*, 72, 293–329.
- Park, V., & Datnow, A. (2008). Collaborative assistance in a highly prescribed whole school reform model: The case of Success for All. *Peabody Journal of Education*, 83(3), 400–422.
- Park, V., & Datnow, A. (2009). Co-constructing distributed leadership: District and school connections in data-driven decision making. *School Leadership and Management*, 29(5), 477–494.
- Patton, M.Q. (2002). *Qualitative evaluation and research methods*. Thousand Oaks, CA: Sage.
- Porter, K. E., & Snipes, J. C. (2006). *The challenge of supporting change: Elementary student achievement and the Bay Area School Reform Collaborative’s focal strategy*. New York, NY: MDR.
- Quint, J., Sepanik, S., & Smith, J. (2008). *Using student data to improve teaching and learning: Findings from an evaluation of the Formative Assessments of Student Thinking in Reading (FAST-R) program in Boston elementary schools*. New York, NY: MDR.
- Reynolds, D., Stringfield, S., & Schaffer, E. (2006). The High Reliability Schools project: Some preliminary results and analyses. In J. Chrispeels & A. Harris (Eds.), *School improvement: International perspectives* (pp. 56–76). London, England: Routledge.
- Roehrig, A. D., Duggar, S. W., Moats, L., Glover, M., & Mincey, B. (2008). When teachers work to use progress monitoring data to inform literacy instruction: Identifying potential supports and challenges. *Remedial and Special Education*, 29(6), 364–382.
- Smircich, L. (1985). Is the concept of culture a paradigm for understanding organizations and ourselves? In P. J. Frost, L. F. Moore, M. R. Louis, C. C. Lundberg, & J. Martin (Eds.), *Organizational culture* (pp. 55–72). Beverly Hills, CA: Sage.
- Slavit, D., & Nelson, T. H. (2010). Collaborative teacher inquiry as a tool for building theory on the development and use of rich mathematical tasks. *Journal of Mathematics Teacher Education*, 13, 201–221.
- Spillane, J. P. (2000). Cognition and policy implementation: District policymakers and the reform of mathematics education. *Cognition and Instruction*, 18(2), 141–179.
- Spillane, J. P., & Miele, D. B. (2007). Evidence in practice: A framing of the terrain. In P. A. Moss (Ed.), *Evidence and decision making* (pp. 46–73). Malden, MA: Blackwell.
- Stringfield, S., Reynolds, D., & Schaffer, E. C. (2008). Improving secondary students’ academic achievement through a focus on reform reliability: 4- and 9-year findings from the High Reliability Schools project. *School Effectiveness and School Improvement*, 19(4) 409–428.

- Supovitz, J. A. (2006). *The case for district-based reform: Leading, building, and sustaining school improvement*. Cambridge, MA: Harvard Education Press.
- Supovitz, J. A. (2012). Getting at student understanding—The key to teachers' use of test data. *Teachers College Record*, 114(11)
- Supovitz, J., & Klein, V. (2003). *Mapping a course for improved student learning: How innovative schools use student performance data to guide improvement*. Philadelphia, PA: Consortium for Policy Research in Education.
- Supovitz, J., & Weathers, J. (2004). *Dashboard lights: Monitoring implementation of district instructional reform strategies*. Philadelphia, PA: Consortium for Policy Research in Education.
- Sutherland, S. (2004). Creating a culture of data use for continuous improvement: A case study of an Edison Project school. *American Journal of Evaluation*, 25(3), 277–293.
- U.S. Department of Education. (2009). *American Recovery and Reinvestment Act of 2009: Title I, Part A Funds for Grants to Local Education Agencies*. Retrieved from <http://www.ed.gov/print/policy/gen/leg/recovery/factsheet/title-i.html>
- Weiss, C. H. (1980). Knowledge creep and decision accretion. *Knowledge: Creation, Diffusion, Utilization*, 1(3), 381–404.
- Wohlstetter, P., Datnow, A., & Park, V. (2008). Creating a system for data-driven decision-making: Applying the principal-agent framework. *School Effectiveness and School Improvement Journal*, 19(3), 239–259.
- Yang, M., Goldstein, H., Rath, T., & Hill, N. (1999) The use of assessment data for school improvement purposes. *Oxford Review of Education*, 25(4), 469–483.
- Young, V. M., & Kim, D. H. (2010). Using assessments for instructional improvement: A literature review. *Education Policy Analysis Archives*, 18(19), 1–36.

Appendix: Description of Literature Reviewed

		Intervention Domains of Action	Study Sample	Study Period (yrs)	Framework	Design & Data Collection Methods	Methods Documented	Triangulation	Findings Supported by Evidence	Methods Concerns/ Strength of Sample(s)	Discuss Reliability, Validity, Credibility	Participant Observer
School System												
2 Districts	Goertz et al., 2009	HS-PD, C, T, N; Tech; D; E	2d, 9s, 45p	2	CF	CS: s, i, o, d	✓	✓	✓			
Standards-Based Dists	Massell, 2001	HS-PD, T, N; A; E	23d, 33s	3		CS: i, s	✓	✓	✓	few details		
	Massell & Goertz, 2002	HS-PD, T, TE, N; D; A; E	23d, 57s	3		CS: i, o	✓	✓	✓	few details		
Demo Dists	Marsh, 2007	HS-T, N; E	2d	3	Th	CS: i, o, d, s	✓	✓	✓	low RRs		
Duvall	Supovitz, 2006	HS-PD, T, TE, N; Tech; D	1d, 10s	7	TOA	CS: i, s, d	✓	✓	✓	good RR		y
	Supovitz & Weathers, 2004	HS-PD, T, TE, N; Tech; D	1d	1		CS: i, s, d	✓	✓	✓	good RR		y
Edison Schools	Marsh et al., 2008	HS-PD, T, N; Tech; D; A; E	1d, 23s	5	TOA	CS: i, o, fg, o	✓	✓	✓		✓	
	Sutherland, 2004	HS-PD, T, N; Tech; D; A; E	1s	2	CF	CS: i, o	✓	✓	✓			
MDRC	Snipes et al., 2003	HS-PD; Tech	6d*	1		CS: i, fg, d	✓	✓	✓	retrospect		
Milwauk.	Moody & Dede, 2008	HS-PD, T; Tech; D; E	1d	2	CF	CS: i, o, d	✓		✓	few details		
43 Districts	Anderson et al., 2010	HS-PD, C, E	43d, 180s	3	CF	CS: s, i	✓	✓	✓	no RRs	✓	
National D	Means et al., 2010	HS-PD, C, T, TE; Te; D; A	529d/ 12d	2	CF	CR & CS: i, s	✓	✓	✓	good RRs		
USC Districts	Park & Datnow, 2009	HS-PD, T, TE; A; E; Tech	4d, 8s	1	CF	CS: i, o, d	✓	✓	✓			
	Wohlstetter et al., 2008	HS-PD, T, TE; A; E; Tech	4d, 8s	1	Th	CS: i, o, d	✓	✓	✓			

UK LEA	Demie, 2003	HS-PD, T, TE	1d	?		CS: s, i	✓		✓	few details		y
UK LEA	Yang et al., 1999	HST	1d	?		CS				no details		y
Intermediary or School Reform Network												
BASRC	Copland, 2003	HS-PD, C, T, N	1io, 18s	1	TOA	s, i, o	✓	✓	✓	no RRs		
	Jaquith & McLaughlin, 2009	HS-PD, C, T, N	1io, 4d	10	TOA	s, i, fg, o	✓		✓	few details		
	Porter & Snipes, 2006	HS-PD, C, T, N	1io, 10d*	2		CS&QE: i, fg, s, ach	✓	✓	✓	no RRs	✓	
EASTR	Quint et al., 2008	HS-C; Tech; D	1d, 57s*	1		QE & s, ach	✓		✓	low RRs		
Gatehouse	Bond et al., 2001	HS-PD, T, TE	1io, 26s*	3	CF	RCT&i, fg, d	✓	✓	✓			
Getting Results	Gallimore et al., 2009	HS-PD, C, T, TE; D	1d; 15s*	5	TOA	QE&CS: fg, i	✓	✓	✓			✓
	McDougall et al., 2007	HS-PD, C, T, TE; D	1d; 15s*	1		QE&CS: i, o, fg, d	✓	✓	✓			✓
High Reliability	Reynolds et al., 2006	HS-PD, T, N	3d, 25s	5		s, d, i, ach	✓		✓	no RRs		y
	Stringfield et al., 2008	HS-PD, T, N	1d, 12s	9		L & CS: i, s, o, ach	✓	✓	✓	few details		y
IFL	Kerr et al., 2006	HS-PD, C, T, TE, N; E	1io, 3d	2	TOA	CS: s, i, d	✓	✓	✓	low RR		
	Honig & Ikemoto, 2008	HS-PD, C, T, TE, N; E	1io, 8d	6	Th	CS: i, fg, o, d	✓	✓	✓			✓
Merck	Corcoran & Lawrence, 2003	HS-PD, TE; D	1io, 4d	10		CS: i, s, o, d	✓		✓	few details		
NCREST	Ancess et al., 2007	HS-PD, T	1io	>1		CS			✓			
Oakland	Honig, 2004	HS-TE, N	4io, 1d	2	Th	CS: o, i, d	✓	✓	✓			
P D Ref.	Coburn, 2010	HS-PD, TE, N	1io, 1d	3	Th	CS: i, o, d	✓	✓	✓			
Success For All	Park & Datnow, 2008	HS-PD, C, T; Tech	1io, 2s	1		CS: i, fg, o, d	✓	✓	✓			
	Datnow & Castellano, 2001	HS-PD, C, T; Tech	1io, 6s	3		CS: i, fg	✓	✓	✓			

Multi- or Cross-Site Professional Development												
Florida Coaching	Marsh et al., 2010	HS-C	8d, 113s	1	TOA	CR & CS: s, i, fg, d, o	✓	✓	✓	mixed RRs	✓	
	Roehrig et al., 2008	HS-C	30p	1		s, i	✓	✓	✓	good RR	✓	
Global PD	Huffman & Kalnin, 2003	HS-PD, T	8 grp, 29p	1		s, fg	✓	✓	✓	good RR		
Portfolios	Gearhart & Osmundson, 2009	HS-PD, N, T	23p	1.5	CF	s, fg, d	✓	✓	✓	good RRs		
PRiSSM	Nelson & Slavit, 2007	HS-PD, T	5 groups	1	Th	CS: o, d, fg	✓	✓	✓			y
	Slavit & Nelson, 2010	HS-PD, T	1 group	1	Th	CS: o, d, i	✓	✓	✓		✓	y
SFC	Denton et al., 2007	HS-C, T	34p*	1		fg, d	✓		✓	small n		
Workshop	Murnane et al., 2005	HS-PD, T, N; Tech	32p	1		s, d, o	✓	✓	✓	good RR		y

Note. Domains of action: HS (human support) – PD (professional development), C (coaching), T (tools), TE (technical expertise), N (networking/brokering); Tech (technology support); D (data production); A (accountability/incentives); E (expectations/culture). Sample: d (district/local education agency/education management organization), s (school), io (intermediary organization), p (participants); * indicates inclusion of comparison group. Framework: TOA (theory of action), CF (conceptual framework), Th (theory). Design & Methods: CR (cross-sectional); CS (case study); L (longitudinal); RCT (randomized controlled experiment); QE (quasi-experimental); s (survey), i (interviews), fg (focus group), d (documents), o (observations), ach (student achievement data). Darker, larger check marks (✓) indicate stronger evidence supporting the category. Strength of Sample: RR (response rate), retrospect (retrospective case).

JULIE A. MARSH is an associate professor at the Rossier School of Education at the University of Southern California. She specializes in research on policy implementation, educational reform, and accountability. Her research blends perspectives in education, sociology, and political science. Recent or current research includes studies of the Los Angeles Unified School District’s Investing in Innovation (i3) grant to support the turnaround of low-performing schools; New York City’s Schoolwide Performance Bonus Program; and efforts to build capacity for data use. Dr. Marsh has authored, coauthored, and coedited numerous publications, including “How Instructional Coaches Support Data-Driven Decision Making: Policy Implementation and Effects in Florida Middle Schools,” *Educational Policy*, 24(6), 872–907; “Assistance and Accountability in Externally Managed Schools: The Case of Edison Schools, Inc.,” *Peabody Journal of Education*, 2008, 83(3), 423–458; “Cutting

Through the ‘Data Driven’ Mantra: Different Conceptions of Data-Driven Decision-Making,” *Evidence and Decision Making: National Society for the Study of Education Yearbook 2007*, 106(1); and *Democratic Dilemmas: Joint Work, Education Politics, and Community* (SUNY Press, 2007).