URBAN DATA STUDY

Using Data to Improve Instruction in the Great City Schools:

Key Dimensions of Practice

December 2010

Authors

Jessica Heppen

Ann-Marie Faria

Kerri Thomsen

Katherine Sawyer

Monika Townsend

Melissa Kutner

Suzanne Stachel

American Institutes for Research

Sharon Lewis

Michael Casserly

Council of the Great City Schools

Contents

Using Data to Improve Instruction in the Great City Schools: Key Dimensions of Practice

Introduction

Recent years have seen increased interest in data-driven decision making in education; that is, using various types of data, particularly quantitative assessment data, to inform a range of decisions in schools and classrooms (Marsh, Pane, & Hamilton, 2006). In addition to being a natural result of the increased availability of quantitative data brought about by accountability reforms, the increased emphasis on using data is facilitated by the belief that assessment and other data can be an important lever for improved teaching and learning. Many districts have invested in benchmark assessments and other mechanisms for gathering additional information about student performance prior to the administration of end-of-year accountability tests. At the same time, districts, states, and schools have invested resources in tools designed to provide teachers, principals, and other key actors with ready access to (and analysis of) information regarding student performance.

Of particular interest in this area is the development of interim assessments, administered at regular intervals throughout the academic year, which are designed to predict student performance on end-of-year accountability tests. As we review in this report, there is a growing body of research on this subject. Previous studies have examined the common features at apparently high-performing schools and districts and have found that "data driven" instruction and decision making are common characteristics in some of these organizations (e.g., Marshall, 2006). They have also examined the implementation of "data practices" in districts and schools that are purportedly making strides in data-driven decision making and instruction or that have undertaken significant initiatives in this area (e.g., Datnow, Park, & Wohlsetter, 2007; Snipes, Doolittle, & Herlihy, 2002).

However, the field has yet to produce reliable evidence regarding the relationship between particular uses of data on the one hand and teacher or school effectiveness at raising student achievement on the other. Relatively few studies have focused on the teacher or school level, attempted to develop specific measures of data use practice, or estimated the relationship between these specific practices and student achievement.

In fall 2008, the Council of the Great City Schools and the American Institutes for Research launched a project funded by the Bill & Melinda Gates Foundation focused on understanding the use of interim assessment data as a lever for instructional improvement. The overarching goal of this study is to establish principles of "best practice" in using interim assessment data to improve instruction and target support.

In this preliminary report, we first review the literature on using data from interim assessments and put forth a Theory of Action that undergirds our investigation. The theory of action identifies a set of Key Dimensions of data use practice and hypothesizes that supporting conditions in states, districts, and schools can facilitate effective classroom-level use of data to respond to students' instructional needs. In the final section of this report, we provide an initial empirical test of this theory of action, using survey data from more than 500 teachers across four urban districts collected during the 2009–2010 school year. The focus in this preliminary report is on the predictors of whether teachers use interim assessment data to change their instruction. Future reports will extend the findings to examine the data-use practices that predict improvements in student achievement.

Literature Review on Using Interim Assessment Data

Overview

The continued advancement of technology and the growing pressure for schools to be data driven have resulted in substantial new funding and research on educational data systems (Hamilton, 2005). Test scores have been used for some time to make instructional decisions, but the availability and use of data to inform such decisions have not always been systematic (Abelman, Elmore, Even, Kenyon, & Marshall, 1999).

Recent efforts to implement systematic assessments at regular intervals during the school year hold promise for higher achievement, and some researchers and practitioners suggest that their use may be critical to school improvement. According to Marshall (2006), many schools that serve disadvantaged students who "beat the odds" academically analyze their interim assessment data as part of their overall strategy for improving achievement. Indeed, studies that have examined common features of high-performing schools and districts have found that data-driven instruction and decision making are common features in many of these organizations (Datnow, Park, & Wohlsetter, 2007; Snipes et al., 2002).

With the U.S. Department of Education's desire to close achievement gaps through data use, new policies have been implemented to promote data use in schools and classrooms. The American Recovery and Reinvestment Act of 2009 called on states, districts, and schools to develop longitudinal data systems to increase their capacity to support students' strengths and identify their weaknesses. This legislation sends a strong message about the importance of using data to inform educational practices and will inform a dialogue among multiple stakeholders on how data should and can be used in the future to improve public education. Other initiatives, such as the Data Quality Campaign (DQC), have also focused attention and resources on building state longitudinal databases that house student-level information for use by stakeholders at all levels (DQC, 2009). Access to student data is clearly growing; for example, according to a nationally representative survey, teacher access to student data systems grew from 48 percent in 2005 to 74 percent in 2007 (U.S. Department of Education, 2009).

However, despite the recent attention and investments at national, state, and local levels, researchers and practitioners have not reached consensus on what being data driven actually means in practice. Moreover, little evidence connects specific data uses to changes in teaching and actual improvements in student outcomes. In the first section of this report, we summarize the research on using data, particularly interim assessment data, to guide instructional decisions and improve student outcomes in the context of a set of Key Dimensions of Practice defining the current thinking around the best strategies for using interim assessments to support improved student achievement. But before we present our review of the literature under the Key Dimensions of Practice section, we address the types of assessments included in our review.

Background on Assessment Types

Relevant literature on the use of regular, periodic assessment data to inform instruction includes various types of assessments and assessment strategies. Our review included research on a range of types of periodic assessments that are administered with regular frequency throughout the school year. In addition to studies of interim (or benchmark) assessments, our review of the literature included studies of formative assessment, progress monitoring, and curriculum-based measurement. Not all these assessments have agreed-on definitions, and they are not mutually exclusive. The goal here is not to create clear delineations between these assessment types but to acknowledge their commonalities in providing information that educators can use to inform and adjust classroom instruction. Indeed, educators may be best able to respond to data if they have access to information about student strengths and weaknesses that derives from a variety of assessment types (Hamilton et al., 2009).

Informal Formative Assessment, Progress Monitoring, and Curriculum-Based Measurement

Formative assessment is an ongoing process in which classroom teachers assess students' knowledge and understanding with activity-embedded, brief, small-scale tasks that are linked directly to the current curriculum topic. Formative assessments are not always standardized across schools, classrooms, or even students; therefore, aggregating formative assessment data is not typically done or useful (Perie, Marion, Gong, & Wurtzel, 2007). A number of studies on formative assessment suggest that classroom-embedded student assessments can be used by teachers to elicit achievement gains (Black, Harrison, Marshall, Lee, & Wiliam, 2002; Brookhart, 2001; Hayward, Priestley, & Young, 2004; Heritage, 2007; Research for Action, 2009; Shepard, 2005). In a seminal piece on the topic, Black and Wiliam (1998a) conducted an extensive review of the literature on formative assessments and concluded that formative assessments can increase student achievement. In a related piece, Black and Wiliam (1998b) suggested that formative assessments are effective because by definition they use evidence to directly inform teaching practices to meet students' learning needs, unlike summative and other assessments. Drawing from eight studies of formative assessment, the authors concluded that increases in formative assessment practices lead to learning gains. None of the eight studies showed that increases in formative assessment practices negatively affected student achievement. However, the magnitude of the increases in student learning varied according to *how* teachers responded to the information that formative assessments provided and how they used the assessment information to provide feedback to students about their progress. Based on this finding, Black and William called for further research, specifically on how teachers can best use assessment feedback to improve student learning.

Curriculum-based measurement (CBM) provides a useful example of student-level progress monitoring. In the special education literature, several studies on the effects of CBM on learning outcomes of students with disabilities provided early insight into data-driven teacher practices that positively affect student achievement (e.g., Fuchs, Deno, & Mirkin, 1984; Deno, 1985; Deno & Fuchs, 1987). CBM specifies procedures for measuring student proficiency within curricular goals and basic skills. Some key uses of CBM are screening students for special services, developing and monitoring instructional programs, and evaluating program efficacy (Fuchs &

Fuchs, 1990). In a study by Fuchs and colleagues (1984), students with disabilities were randomly assigned either to CBM treatment or to a more traditional special education evaluation over a period of 18 weeks. The researchers observed teacher pedagogy and concluded that teachers assigned to use CBM were more responsive to their students' needs and achievements than those in more traditional special education settings. Moreover, their students achieved greater outcome gains than their peers in the control group. Similarly, Fuchs and Fuchs (1990) also demonstrated that CBM increased student learning outcomes and that CBM with additional support (in the form of skills analysis) increased student performance more than using CBM alone. This result emphasizes the focus of CBM on providing teachers with instructional guidance, in addition to providing high-quality tools for ongoing assessment, to affect student learning and understanding.

Other studies of CBM report similar findings, demonstrating that CBM increases student achievement (Davis & Fuchs, 1995; Hintze, Daly, & Shapiro, 1998; Marston et al., 2007; Stecker & Fuchs, 2000) and, perhaps more important, when coupled with instruction in effective teaching strategies increases student outcomes more than CBM alone (Fuchs, Fuchs, Hamlett, & Stecker, 1991; Fuchs, Fuchs, & Hamlett, 2007).

The existing evidence about formative assessments in general and their use in specific contexts (e.g., special education) provides a basis for the use of regular, systematic assessment to inform instruction. However, research has yet to clarify whether the widespread use of assessments—commonly known as "interim" or "benchmark" assessments—can produce robust gains in student achievement.

Interim and Benchmark Assessments

Interim assessments are assessments that are administered at regular intervals throughout the school year to help educators gauge student achievement before the annual state exams used to measure Adequate Yearly Progress (AYP; Research for Action, 2009). Interim assessments provide data that can be aggregated or disaggregated to the student, teacher, and school levels and are often designed to predict student performance on end-of-year accountability assessments. Other stated purposes of interim assessments are to provide information to diagnose student strengths and weaknesses and to provide evaluative information about curricula or instructional programs (Perie et al., 2007). Characteristics of interim assessments, also known as benchmark assessments, are that they are administered routinely (e.g., every 6 to 8 weeks) across grade levels in particular content areas (e.g., reading or mathematics) within a school or district. They may be commercially developed, developed by districts or states, or a combination. Some interim assessments are delivered as fixed-form tests, whereas others are delivered as computer-adaptive tests based on large item banks.

Not surprisingly, some evidence suggests that such assessments are not sufficient by themselves to raise student achievement. In one of the more rigorous analyses of the effectiveness of interim assessments to date, Henderson, Petrosino, Guckenberg, and Hamilton (2007, 2009) reported that benchmark assessments used as part of a pilot program in Massachusetts did not yield improvements in student mathematics achievement. In the study, 22 schools used state-developed quarterly administered benchmark assessments. Using a quasi-experimental design, the researchers matched these schools in terms of student population to 44 similar schools. They compared the mathematics achievement scores of eighth-grade students at schools with state benchmark assessments with those of their counterparts in schools without state benchmark assessments. The authors did not observe any statistically significant or substantively important differences between the two groups. However, they note that other interim or benchmark assessments may have been in place at the comparison schools and, perhaps most important, that information about *how* the data from the benchmark assessments were used by educators in the "treatment" schools was outside the scope of the study (Henderson et al., 2007).

Indeed, other research is clear that the ways that interim assessments are implemented and used are key to improving student achievement (Marshall, 2006). The *use* of interim assessment data is the direct focus of the key dimensions of practice and the theory of action that we posit in the current study.

Key Dimensions of Practice

As noted, simply having an assessment system in place is likely not enough to improve student achievement. We offer in this brief a more specific theory of how data practices at multiple levels (district, school, principal, teacher, and student) are related to improved student achievement.

According to the general theory of action implicit in the implementation of periodic, regular assessments, educators can use these data to do the following:

- Better understand the academic needs of individual students, and respond to these needs by targeting instruction, support, and resources accordingly

- Better understand the instructional strengths and weaknesses of individual teachers, and use this information to focus professional development (PD), peer support, and improvement efforts

- Support and facilitate conversations among teachers and instructional leaders regarding strategies for improving instruction

These practices, in turn, are thought to lead to improved and more responsive teaching and therefore yield increased student achievement.

From this broad theoretical perspective, our goal was to articulate a more specific theory of action that supports our current investigation of the relationships among data-use practices and improvements in student achievement over time in large urban districts. Our intention was to ground the specific classroom- and school-level data-use practices that could theoretically improve student achievement in the context of the larger systems in which they occur.

At the outset, we identified four Key Dimensions of practice in data-driven instruction:

1. **Context:** State, district, and school conditions that facilitate a general emphasis on using data, or a "culture of data use"

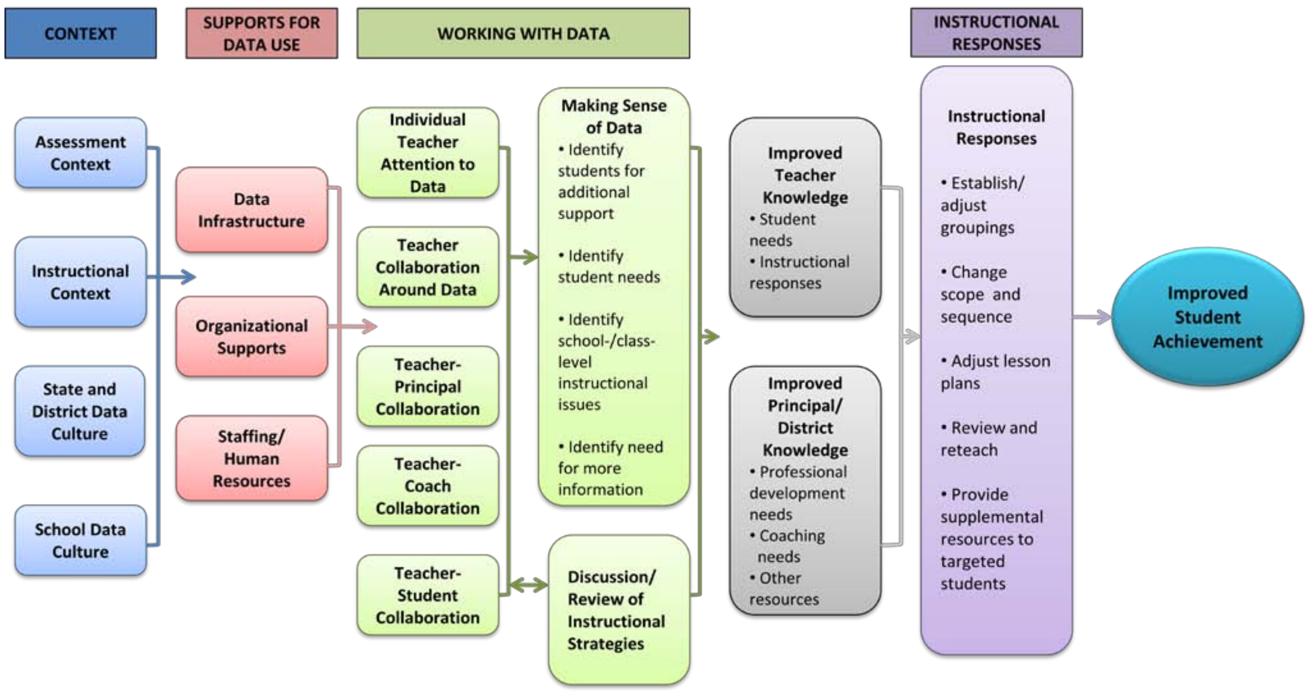
2. **Supports for Data Use:** Concrete factors that facilitate, enable, and support specific uses of data

3. **Working With Data:** The manner in which teachers work both individually and collaboratively to review data and to identify specific ways the data are used to improve teachers' knowledge about student needs

4. **Instructional Responses:** The ways that teachers can respond to the knowledge and information generated by their review of student data, which in theory could lead to improved teaching and learning and higher student achievement outcomes

A theory of action based on these key dimensions is shown in Figure 1.[1]

Figure 1. Using Data From Interim Assessments to Improve Student Achievement



We also posit that to result in Instructional Responses, the ways that educators work with data must result in a change in their knowledge. That is, for a teacher to make an instructional change in the classroom with all or some students, that teacher must have an improved understanding of what the students do or do not know or understand. However, we did not define Knowledge as a Key Dimension of Practice because our focus is on observable, measureable aspects of data use. Nevertheless, we identify Improved Knowledge as a mediating step between Working With Data and Instructional Responses in the theory of action.

In the following sections, we review the literature that informs the theory of action, focusing on the four Key Dimensions. Under each dimension, we define the key elements that are shown as

[1] Although the arrows in the figure attempt to illustrate some of the key causal relationships, the diagram does not attempt to illustrate every important causal relationship. For example, it is very clear that in addition to affecting the outcomes in the column to the right, the elements of practice listed within each of these key dimensions theoretically contribute to and reinforce one another. For example, although "staffing resources" and "organizational supports" both contribute to individual teacher attention to data, "staffing resources" and "organizational supports" should also reinforce one another.

the boxes within the theory of action in Figure 1. Following a review of the relevant literature that informs the key elements, we list a set of important aspects of each key element that provide a framework for the broader study. These aspects were also our specific focus for the development of survey instruments that were used to measure the data use practices represented by the theory of action. The survey study is described in the section that follows the literature review on the Key Dimensions.

1. Context

We first consider the contextual factors that may affect the type of data to which schools have access and how school-level personnel use these data to alter instruction. Key elements of context include the (a) assessment context, (b) instructional context, (c) district-data culture, and (d) school-data culture. Although other contextual elements are theoretically relevant (e.g., the political or economic context), our theory and measurement of data use focus on factors that we hypothesized are most relevant to the use of data in districts, schools, and classrooms.

(a) Assessment Context. Assessment context includes the goals, expectations, and policies related to the development and implementation of interim assessments, including the types of assessments given and their purpose(s). An important aspect of the assessment context also concerns the "quality" (e.g., the validity and reliability) of the interim assessments.

Previous research suggests that the perceived quality of the data is as important as the strength of the data infrastructure and the accessibility and timeliness of the data. The perception that the assessments, the data they yield, or the reports from them are of poor quality can be a clear barrier to use. Many studies have concluded that doubts about the accuracy of data lead to a lack of support for data initiatives and result in decreased data use by teachers (e.g., Feldman & Tung, 2001; Herman & Gribbons, 2001; Herman, Yamashiro, Lefkowitz, & Trusela, 2008; Ingram, Louis, & Schroeder, 2004). In an evaluation of the districtwide use of data in Natrona County, Wyoming, Wayman, Cho, and Johnston (2007) found that concerns about the accuracy of student data were correlated with lower levels of data use. Similarly, Kerr and colleagues (2006) found that the perceived validity of data affected the extent to which teachers used them.

The alignment between the selected assessment and the intended uses of the data is important. Militello, Schweid, and Sireci (2010) emphasize the importance of a close fit between the characteristics of a formative assessment and the intended use of the data by the district and teachers to ensure that the data have a meaningful impact on teachers and students.

Indeed, multiple studies have found that the alignment of the interim assessments with standards, state tests, and pacing guides facilitates data use (Kerr et al., 2006; Marshall, 2008; U.S. Department of Education, 2010). For example, in Philadelphia, researchers found that benchmark assessment results were viewed as highly relevant to teachers' instructional planning because they were aligned to the curriculum and a 6-week instructional cycle. The sixth week of the cycle was designated for remediation and extension of topics that could be designed by teachers on the basis of their review of the benchmark assessment results (Research for Action, 2009). However, it is important to note that in the same study, the authors found that teachers' satisfaction with the benchmarks was not significantly related to student achievement growth in reading or mathematics.

Important aspects of the **Assessment Context:**

- Primary purpose(s) of interim assessments (e.g., to improve instruction, to predict performance on accountability tests)

- Type of assessments given, such as
  - Curriculum embedded
  - External benchmarks
  - Formative versus summative assessments
  - Development and structure, including test construction (e.g., externally versus internally developed), fixed forms versus item banks, cumulative versus unit-based assessments, and length and structure of assessments (e.g., number of items per objective)

- Quality of assessments
  - Reliability and validity of the assessments
  - Sufficient and appropriate scaling
  - Alignment with state standards
  - Alignment with curricula and pacing guides
  - Alignment with state assessments

(b) Instructional Context. The instructional context includes the curricular and instructional environment in which teachers and principals collect and use data. The uniformity, focus, and history of the instructional program all have the potential to affect how data are used at the district, school, and classroom levels. As is the case with assessment context, instructional context is primarily meant to capture issues at the *state and district levels* that shape school and classroom data activities.

A key aspect of instructional context is the degree of flexibility in the curriculum and pacing schedule. The literature suggests that districts and schools must be flexible in their curriculum pacing to allow teachers time to alter instruction based on assessment results (Clune & White, 2008; Datnow et al., 2008; David, 2008). Marsh and colleagues (2006) suggest that curriculum pacing pressures, especially in the presence of regimented programs with pacing plans, are an obstacle to data use by teachers. Even if pacing pressures were more perceived than real, teachers often follow the pacing plans instead of adjusting their instruction on the basis of the results of their data analyses (Marsh et al., 2006).

Important aspects of the **Instructional Context:**

- Presence and implementation of districtwide curricula and pacing guides, including
  - Flexibility of curricula
  - Flexibility and speed of pacing (including time to reteach)

- Centralized *versus* site-based decision-making, including instructional and curricular decisions, staffing/human resources decisions, decisions about professional development

- Presence and implementation of systemwide or schoolwide Response to Intervention strategies

- Presence and implementation of other districtwide or statewide instructional initiatives

- Accountability context, including district and school Adequate Yearly Progress history

(c) State and District Data Culture. State and district data culture includes attitudes, direction, and support at state and district levels regarding the use of data in general and interim assessments in particular. We hypothesized that the degree and nature of support for data use, as well as the direction of district and state policy in this area, can affect the manner in which and the extent to which data are employed at the building level.

Marsh and colleagues (2006) found that teachers use data more frequently in school systems whose administrators had committed to data-driven decision making and had a clear vision about data use at the school level. These school systems were also characterized by openness and a sense of collaboration around data use, in contrast to school systems in which data analysis was seen as an individual activity.

Clearly articulated and communicated goals for district data use are also important. Studies have found that a barrier to teacher data use is the perception—real or imagined—that teachers are going to be blamed for the poor performance of their students (Clune & White, 2008; Ingram et al., 2004; Kerr et al., 2006; Marshall, 2008).

A study by Anderson, Leithwood, and Strauss (2010) reinforced the importance of district leadership in setting expectations for data use. They found that district leadership had more power to affect some of the typical barriers to data use, such as accessibility, timeliness, quality, and capacity for use, than did school principals.

Important aspects of **State and District Data Culture:**

- Explicit support for the use of data (i.e., as an explicitly stated state or district priority)

- Clear goals for the use of data across the system

- Clearly articulated plan for implementing processes and procedures to support and encourage data use

- Participation in discussions about data

- Integration of data into state and district reviews, evaluations, and goal setting

- State- and district-level perceptions of validity, relevance, and quality of assessments

(d) School-Level Data Culture. This key element of Context is related to goals, norms, expectations, processes, attitudes, and school-level leadership for the use of interim assessment data at the building level.

Studies such as those of Henderson and colleagues (2007, 2009) described above suggest that having benchmark assessments alone is not sufficient to yield increases in student achievement. Research for Action (2009) found similar results in a study of the implementation of Philadelphia's benchmark assessments. Although teachers expressed satisfaction with the benchmarks and their alignment with the core curriculum, the pacing plan, and the instructional cycle, their satisfaction alone was not predictive of growth in student achievement in reading and mathematics. The key supporting factors that appeared to facilitate the link between use of the benchmarks and academic progress were instructional leadership and collective responsibility for data.

Other research also suggests that leadership is a key factor in the successful use of data. Kerr and colleagues (2006) found greater data use in schools that had created data-driven cultures through strong school and district leadership. Murnane, Sharkey, and Boudett (2005) found that teachers' own use of data depends largely on the amount of principal support for data use. Many other studies also highlight this point (e.g., Feldman & Tung, 2001; Lachat & Smith, 2005; Mason, 2002; Marsh et al., 2006). Anderson and colleagues (2010), however, found that although principals play a key role in influencing data use in their schools, the majority often do not act to change the specific conditions that affect data use that are under their control.

Although the principal has often been identified as the leader responsible for several important data supports, case studies by the U.S. Department of Education (2009) suggest that other individuals, including coaches and lead teachers, may also provide important leadership support for data use. Indeed, schools with higher levels of data use often had more widespread data expertise than typical schools because they did not confine the expertise to the principal or a lead teacher (Anderson et al., 2010).

Important aspects of **School-level Data Culture:**

- Shared goals for the use of data and interim assessments
- Clearly articulated process for the implementation of the data systems
- Presence of an action plan or clear strategy for using data
- Awareness of the action plan
- Level of buy-in for data strategy
- Connection between efforts to use data and overall school improvement plans
- Perceptions of the quality of the usefulness and usability of the data and reports
- Perceptions of data-use training, including buy-in and perceived usefulness of training, expertise of trainers and coaches, and appropriateness of training and coaching activities
- Perceptions of accountability for using data (i.e., that using data is part of job expectations)

- Shared norms on data use and on collaboration around data

- Support and guidance from school- and district-level leadership, including
  - Participation in planned meetings to discuss data
  - Interest in and availability for additional conversations about data, including the extent to which the principal and/or district leadership
    - is seen as a resource for teachers as they use data
    - reinforces expectations and offers support for the use of data
    - focuses on data in the overall school improvement process

2. Supports for Data Use

This Key Dimension involves the specific elements of practice related to logistical and operational support for the use of data, including the infrastructure, organizational, time allocation, and personnel resources necessary to support the uses of interim assessment data to guide and improve instruction. In particular, the key elements in Supports for Data Use are (a) data infrastructure, (b) organizational supports, and (c) staffing and human resources.

The concrete supports that districts and schools can provide to enhance data use are important. In an article summarizing their two studies of benchmark data use in Philadelphia, Bulkley, Christman, Goertz, and Lawrence (2010) assert that benchmarks can serve an instructional purpose, but critical to such use are the supports provided by the district, including data systems, useful reports, time for reflection and collaboration, and professional development. Marsh and colleagues (2006), reflecting on conclusions drawn from four studies conducted by the RAND Corporation, also suggest that concrete support for data use is critical to encouraging teachers to use data. The concrete supports they emphasize include various infrastructure supports such as data access, timeliness of the data reports, and adequate time for teachers and principals to review and discuss data.

(a) Data Infrastructure. This key element includes the infrastructure for disseminating, accessing, analyzing, and using data. This dimension is related to the amount of investment and support that exist at the district level, but it is focused on the tools and resources that are available at the school level. It includes both technology-related resources and the content of the data and reporting system itself. This dimension consists of two primary elements: the infrastructure for access to and dissemination of data and the content or capacity of the reports and data system.

Several studies (Datnow et al., 2008; Kerr et al., 2006; Murnane et al., 2005; Wayman & Stringfield, 2006) have emphasized the importance of system-level infrastructure support. A U.S. Department of Education study (2009) found that data systems are often of limited use to teachers for instruction because of limitations in the data, user interface, or system tools. For example, only slightly more than half of teachers with access to a data system also reported having access to their students' benchmark or diagnostic test performance. However, 79 percent of *districts* report having an assessment system that analyzes benchmark data. Therefore, there appears to be a disconnect between district-level infrastructure and teachers' perceptions of access to the infrastructure.

The quality of the data infrastructure also seems to affect levels of teacher data use. Kerr and colleagues (2006) found that schools demonstrating greater data use had better data infrastructure systems that included timely reporting of results, online access to data, and an interface that allowed teachers to manipulate data and run specialized reports.

Other studies concur that timeliness and accessibility of data are particularly important. For example, Clune and White (2008) concluded that out-of-date data significantly impede the ability of teachers to modify their instruction. Teachers received data approximately 2 weeks after assessments were administered, which is a time frame that teachers regarded as too late to modify instruction in the current school year. Similarly, in a study of data use in five urban high schools, interviews with teachers and principals revealed that those who had access to timely data were more likely to use them and were more successful at integrating results from their analyses into classroom practice (Lachat & Smith, 2005).

Important aspects of **Data Infrastructure:**

- Data Access and Dissemination
  - Type of access: availability of direct access into the data system at the school level, provision of district-generated reports in such formats as electronic, paper, or online
  - Level of access by role (e.g., district, school principal, teacher, student, parent)
  - Availability of computer resources
  - Ease of access and use
  - Frequency and timeliness of reports
  - Ability to manipulate data (e.g., disaggregated by item types or student subgroups)

- Content and Capacity of Reports and Data Systems
  - Identification of specific student needs, including "at-risk" students, "bubble" students, and areas of student strengths, weaknesses, and misconceptions
  - Identification of classroom-level needs and challenges
  - Identification of school-level needs and challenges
  - Data disaggregation (i.e., student performance by subgroup, content standards, item types)

(b) Organizational Supports. This key element refers to logistical and operational supports for data use, including scheduling and allocating time for review and discussion of interim assessment data and their implications for instruction. Although the presence or prevalence of these supports may be a function of the data culture within a school or district, this dimension is focused on concrete supports that exist apart from norms, expectations, and other "soft" supports.

A key supporting factor for the use of data in schools is the allocation of time for teachers to work independently and collaboratively with student data. That is, the lack of structured time to learn how to use data, review data, and reflect on instructional responses is often cited as a barrier to effective data use. A report by the U.S. Department of Education (2009) noted that the majority of teachers who use student data report doing so on their own *and* with colleagues in their department or grade level, in grade-level team meetings that are sometimes facilitated by a coach.

Other studies have also found that data use is more likely to occur when districts provide structured time to allow teachers to learn how to use data collaboratively (Young, 2006). In addition, tools such as assessment results linked to model lesson plans, frameworks, and curriculum guides can be developed within the data system to help teachers interpret data and respond instructionally; however, these are not common—even in districts known to be "high data users" (U.S. Department of Education, 2009).

Important aspects of **Organizational Supports:**

- Common planning time and regular meetings to discuss data
- Time allocated and available for teachers to access and review data during the school day
- Principal participation in data-focused meetings
- Time allocated for one-on-one meetings between teacher and principal or a "data leader" to review and discuss classroom or student assessment data
- Monitoring and implementation support: district- and school-level oversight, evaluation, reporting requirements, tools or protocols for reviewing and understanding data, templates for action plans, and other monitoring and/or assistance in creating and maintaining a data-use process

(c) Staffing and Human Resources. This key element refers to the human resources and training that affect a school's capacity to use data to improve instruction. It includes staff positions, capacity of staff to use data, and professional development available to support data use.

Lack of staff capacity and lack of training in assessment and data analysis have been reported as important obstacles to teacher data use in numerous studies (Heritage, 2007; Heritage & Bailey, 2006; Herman et al., 2008; Ingram et al., 2004; Lachat & Smith, 2005; Sharkey & Murnane, 2006; U.S. Department of Education, 2009; Wayman et al., 2007). Herman and colleagues (2008) found that teachers are typically not trained in assessment and are not taught the content and pedagogical knowledge required to interpret student performance results and make instructional changes.

Professional development in the use of data provided by instructional coaches or other data facilitators can increase the likelihood that teachers will use data—that is, several studies suggest that trained teachers are also more apt to modify their teaching practices appropriately on the basis of the knowledge they have gained from assessment data (Henke, 2005; Marsh et al., 2006; Mason, 2002). Young (2006) conducted observations and interviews with district administrators, school principals, and teachers about data use and concluded that effective data use was more likely to occur when districts modeled data use for their teachers.

Based on the 2006-2007 National Educational Technology Trends Study (NETTS) teacher survey, 39 percent of teachers self-reported that the training they received about data-driven decision making prepared them to use data to improve student achievement (U.S. Department of Education, 2009).

In addition, school districts with high levels of data-driven decision making tend to offer district-funded, school-based data coaches to support teachers' data use. The role of the data coach varies, but typical responsibilities include helping teachers examine and interpret data and connect results to instructional strategies (U.S. Department of Education, 2009).

Important aspects of **Staffing and Human Resources:**

- Personnel resources: The availability of staff resources dedicated to data use:
  - Dedicated school-level staff to support access to and analysis of data (such staff could be hired by the school or district or be a teacher, specialist, coach, or other staff person)
  - District-level staff available to support work with data (e.g., data coaches, instructional support staff who work on data use)
  - The functions and goals of data support staff (e.g., instructional versus accountability versus progress monitoring)
  - Availability of data support staff: percentage of time dedicated to providing data use support versus other responsibilities

- Staff capacity: The attributes, prior training, and expertise of staff, which are meant to reflect the capacity of staff to access, interpret, and respond to the assessment data:
  - Assessment literacy
  - Awareness of available data systems
  - Experience using data systems
  - Prior training[2]

- Professional development and training: The availability, participation, quality, and focus of training concentrated in order to use data effectively, which is, in theory, a key feature of effective data use for instructional improvement. This dimension is defined in a broad sense to capture training administered in workshops or summer institutes as well as "real time" supports such as coaching or other classroom-based supports.
  - Access to and availability of data-use coaching, including
    - presence of coaches
    - content of coaching
    - type of coaching: mandatory versus voluntary
    - duration and amount of coaching
    - expertise and experience of coaches
  - Level of participation in data-use training activities
  - Presence of ongoing evaluation of the success and effectiveness of training activities

[2] Although staff attitudes toward using data for instruction could be considered an element of staff capacity, we are incorporating staff attitudes toward data use into the school data culture category.

3. Working With Data

This Key Dimension includes the nature and extent of school activity focused on data and the specific ways in which classroom-, school-, and district-level staff review and understand interim assessment data, interact with one another regarding assessment data, and use these data to inform their knowledge of student needs and decision making regarding instructional strategies. In other words, what do they do with the data they receive? How do they interpret the data and use the data to inform what they know about student needs? This dimension has seven key elements: (a) individual teacher attention to data, (b) teacher collaboration around data, (c) teacher-principal collaboration, (d) teacher-coach collaboration, (e) teacher-student collaboration, and (f) "making sense of data," that is, specific ways to review assessment data to understand student performance.

We hypothesize that these elements of practice are mutually reinforcing. Although individual attention to data and collaboration around data are a function of the supports for data use, they also are a function of one another. For example, teacher collaboration around data supports individual teacher attention to data, and vice versa. Specific ways to make sense of data are facilitated by the time spent in collaboration around data (and by individual teacher attention to data), and they are a function of the extent to which collaboration around and individual attention to data emphasizes these particular activities. Collaboration around data, individual teacher attention to data, and specific ways to make sense of data all lead to discussion and review of instructional strategies and responses. This entire process leads to improved teacher knowledge of student needs and available instructional responses. This in turn leads to specific instructional responses (our fourth category of practice). We define each of the seven elements of Working With Data after a review of the literature about the ways that educators review student assessment data.

Research is clear that simply having interim or benchmark assessments in place is not enough—knowing how to use data to inform instructional practice is essential to improving student achievement. Specifically, to be most effective, educators must first use data to identify the problem and then must move beyond identifying the problem to identifying the reasons behind the problem and how to take appropriate actions to solve it (Anderson et al., 2010).

Multiple research studies have noted some commonly cited uses of student assessment data, including identifying individual or groups of students with particular needs (Henke, 2005; Love, 2004; Niemi, Vallone, Wang, & Griffin, 2007); identifying "bubble" or "at risk" students whose scores fall within particular ranges (Marsh et al., 2006; Long, Rivas, Light, & Mandinach, 2008; Research for Action, 2009; Blanc et al., 2010); and comparing classroom scores with school scores (Niemi et al., 2007).

A study in Philadelphia's public schools revealed that teachers consider a variety of factors when reviewing data (Nabors Ol'ah, Lawrence, & Riggan, 2010). Most teachers in the study began their review by identifying their classes' weak points. They considered their students' results in light of what they knew about their students' background or performance, their district's curriculum or pacing guide, and their perceptions about the difficulty of the material for their students. They used a triage method of focusing efforts on the students or topics deemed most in need of attention.

Henke's (2005) case study of three school districts provides examples of how teachers and principals used data from district-level interim assessments in efforts to improve student achievement outcomes. Seven of eight schools in the Lemon Grove School District were ranked "underachieving" by the California Academic Performance Index. After implementing a new districtwide data system, three of the four schools that received Title I aid were declared High Achieving Title I schools. The author noted that implementation of the data system alone did not increase student achievement. Instead, improvement was attributed to the principals' and teachers' specific and targeted use of data to guide instruction and intervention. For example, in one school the principal analyzed state and district assessment results to identify students with particular weaknesses. The principal found that high percentages of fifth-grade students were performing below proficiency in vocabulary and reading comprehension. This knowledge shaped the school's literacy intervention, which in turn was believed to have led to improvements in student reading achievement.

In other cases, however, similar uses of data did not produce increases in student outcomes, or at best, student progress was unclear. For example, Quint, Sepanik, and Smith (2008) investigated the effects of the Formative Assessments of Student Thinking in Reading (FAST-R) program in the Boston Public Schools (BPS). They examined trends in third- and fourth-grade reading scores on the Massachusetts Comprehensive Assessment System in 21 BPS elementary schools and found that the FAST-R data-use program did not have a statistically significant effect on student achievement. However, the study concluded that the FAST-R program, and in particular its data coaches, seemed to help teachers better understand how their students were performing, including identifying students who were weaker in certain areas of reading.

Similarly, Herman and colleagues (2008) evaluated the effects of the Seattle Public Schools' Comprehensive Value-Added Assessment System on student achievement in 13 schools. The authors found substantial variation in the use of data, and their analysis concluded that data use did not have a significant effect on student outcomes. However, in at least one of four case-study schools, the principal reported that most teachers collaborated frequently to compare student data across classrooms and grade levels to better prepare students for the next grade. The extent to which the lack of effects on student achievement stemmed from variations in data practice, a small sample size, or the lack of effective and widespread data use is unclear, but it appeared that teachers valued the data they had.

In another study of uses of data, Clune and White (2008) examined how data from the quarterly assessments were used in the Providence Public School District. The assessment program was initiated in 2004 and discontinued in 2007. On surveys, teachers reported that they used assessment data to identify students, shape school achievement goals, and modify their instruction.

Studies have found the process of reviewing data to often be collaborative, with teachers consulting with principals, other teachers, or coaches. Data use was found to be a collaborative process in Anderson and colleagues' study of principal data use (2010), involving the principal and teachers in a variety of settings, with the principal more often facilitating teachers' data use than using data themselves. The 2007 National Educational Technology Trends Study found that teacher collaboration around data was almost as common as teachers' individual use of data systems (U.S. Department of Education, 2009). Marsh, McCombs, and Martorell (2009) found a small but statistically significant link between student achievement and the frequency of meetings between coaches and reading teachers to review assessment data.

Involving students in the review of their own data has also been noted as an important aspect of effective data use. Work conducted in the 1990s documented that interventions that incorporated students analyzing their own data combined with feedback from their teachers seemed to improve student outcomes (Phillips, Hamlett, Fuchs, & Fuchs, 1993). Another more recent study that used random assignment found that granting students access to an online-based feedback system that included the students' individual test scores led to increased student achievement (May & Robinson, 2007).

According to recommendations from a recent U.S. Department of Education Practice Guide titled "Using Student Achievement Data to Support Instructional Decision Making," students should be active partners in analyzing their own achievement data. The authors specifically suggest that it is important to explain expectations and assessments to students, including the content and skills that will be assessed (Hamilton et al., 2009). Also, similar to facilitating factors of data use at the teacher and school levels, feedback to students should be timely, clear, and constructive (Black & Wiliam, 1998a; Brunner et al., 2005).

(a) Individual Teacher Attention to Data. This key element includes the extent to which teachers review and interpret interim assessment data to inform their knowledge about student performance.

Important aspects of **Individual Teacher Attention to Data:**

- Frequency with which teachers review interim assessment data and/or reports

- Focus and purpose of review, such as
  - Understanding the data system or reporting system itself
  - Using data to identify and discuss individual student-level issues or needs
  - Using data to identify and understand classroom- or school-level patterns and issues
  - Using data to reflect on instructional challenges and potential solutions
  - Incorporating interim assessment data into (regular) lesson planning

(b) Teacher Collaboration. This key element refers to the amount and nature of teacher collaboration in working with student data. Whereas Organizational Supports described under Supports for Data Use capture the extent to which the school has formally and explicitly allocated time and resources for collaboration, this dimension captures the actual extent of collaboration among teachers.

Important aspects of **Teacher Collaboration:**

- Frequency of teacher meetings about student data

- Level of teacher, principal, and other staff (e.g., district liaisons) participation in meetings about data or involving data discussions

- Focus of data-related teacher meetings, such as
  - Understanding data systems and interpreting assessment results
  - Using data to identify and discuss student-level issues
  - Discussing classroom and school-level patterns in data
  - Using data to reflect on instructional challenges and potential solutions

- Teacher partnering outside of formal meetings, such as one-on-one peer support for interpreting data and developing instructional responses

(c) Teacher-Principal Collaboration. This key element captures the extent and nature of the collaboration between teachers and principals on the use of interim assessment data.

Important aspects of **Teacher-Principal Collaboration:**

- Frequency of teacher-principal meetings about data

- Focus of data-related teacher meetings, such as
  - Recognizing accountability versus support
  - Using data to identify and discuss student-level issues
  - Discussing classroom- and school-level patterns in data
  - using data to reflect on instructional challenges and potential solutions

- Informal teacher-principal interactions about interim assessment data, student results, and student needs

(d) Teacher-Coach Collaboration. This key element captures the extent and nature of the collaboration between teachers and coaches on the use of interim assessment data.

Important aspects of **Teacher-Coach Collaboration:**

- Frequency of teacher-coach meetings and discussions about data

- Focus of data-related teacher-coach meetings

- Focus of data-related teacher meetings, such as
  - Understanding data systems and interpreting assessment results
  - Using data to identify and discuss student-level issues
  - Discussing classroom- and school-level patterns in data
  - Using data to reflect on instructional challenges and potential solutions
  - Working on instructional solutions (e.g., lesson planning, modeling)

(e) Teacher-Student Collaboration. This key element captures the extent and nature of the involvement of students in the review of their own performance on interim assessments.

Important aspects of **Teacher-Student Collaboration:**

- Providing explicit instruction to students on how to use their own achievement data to monitor their progress; motivating students by setting clear, attainable goals; and fostering a sense of control for students over their own learning

- Providing students with a clear understanding of the content and skills that will be assessed

- Providing access to individual test scores for each student with consistent, timely, clear, and constructive feedback from teachers that emphasizes solutions over criticism

- Providing time for reflection on students' own performance

(f) Making Sense of Data. This key element captures the specific ways that classroom-, school-, and district-level staff examine interim assessment data to inform their knowledge about student needs; in other words, the ways that they understand the data they receive—either individually or in collaboration with other teachers, principals, coaches, or students.

Important aspects of **Making Sense of Data:**

- Identification of students for additional support
  - Comparing individual student scores with the performance of a larger group (e.g., class, grade level)
  - Identifying "bubble" students (students below but close to proficiency)
  - Identifying (diagnosing) students with particular needs in foundational skills (e.g., literacy)
  - Identifying students for intervention within the classroom
  - Targeting students for intervention outside the classroom (supplemental or pull-out)

- Identification of specific student needs
  - Identifying students with particular needs in specific concepts
  - Considering groupings of students based on similar patterns or trends over time
  - Reviewing scores on individual items to understand patterns of performance and diagnose areas of misunderstanding (e.g. Item analysis)
  - Reviewing individual scores by content standards and item types
  - Reviewing student growth over time

- Identification of school- and classroom-level instructional issues
  - reviewing average scores to determine class strengths and weaknesses
  - reviewing classroom level scores by content standards, item types, subgroups
  - reviewing student growth over time

- Identification of teachers in need of additional or enhanced PD

- Identification of students or classrooms for instructional interventions

Improved Knowledge

Under Working With Data, we focused on the ways that educators make sense of data (individually and collaboratively). The next Key Dimension, Instructional Responses, focuses on the ways that they make use of that information by changing what they do in the classroom and school. Implicit in the path from making sense of data to implementing instructional responses is a change in **knowledge** among teachers, principals, and district staff. That is, if data use is to be an effective means for improving instruction, it is likely that it must first yield improved teacher knowledge about student needs and principal and district knowledge about teacher and school needs.

Important aspects of **Improved Teacher Knowledge:**

- Awareness and understanding of instructional needs and challenges of individual students
- Awareness and understanding of instructional needs and challenges facing their classrooms as a whole
- Understanding of teachers' own strengths and weaknesses
- Awareness and understanding of strategies and resources for addressing the needs of struggling students
- Awareness and understanding of different strategies for teaching and reteaching specific concepts

Important aspects of **Improved Principal and District Knowledge:**

- Awareness and understanding of instructional needs and challenges facing individual classrooms or teachers and the school as a whole
- Understanding of teachers' (and schools') strengths and weaknesses
- Awareness and understanding of strategies and resources for addressing the needs of teachers and schools

It is important to note that the current study of interim assessment use in urban districts does not directly assess or measure these changes in knowledge. Although we identify knowledge explicitly in the Theory of Action, changes in knowledge are assumed to be implicit between the path from Working With Data to Instructional Responses.

4. Instructional Responses

This Key Dimension captures the specific ways in which classroom-, school-, and district-level staff translate the improved knowledge they glean from reviewing interim assessment results and use it to change classroom-level instruction. This dimension also includes instructional responses (e.g., interventions) implemented at the school and district levels in response to patterns and trends in student assessment data.

As noted under Working With Data, knowing how to use data to inform instructional practice may be essential to improving student achievement. Given a number of studies that have shown that just administering the tests does not appear to yield changes in student achievement, we hypothesize that this process must move beyond using data to diagnose problems to identifying appropriate actions to solve them (see also Anderson et al., 2010).

For example, in the previously described study of Philadelphia's benchmark assessments, Nabors and colleagues (2010) examined the ways that teachers weighed their options for re-covering particular content during the district's "reteaching week." As noted under Working With Data, they used a triage method to identify the topics that were most problematic for students. Teachers in the study most commonly attributed student mistakes to procedural errors; consequently, their reteaching often focused on procedural steps. The Philadelphia assessment results appeared to be of little help in guiding teachers in correcting conceptual errors. Teachers often presented material with which students had struggled in a different way, but the change was not related to an analysis of the assessment items but rather to a belief that being exposed to different methods of teaching was beneficial to students.

Henke (2005) described a school district's initiative to use data in targeted and specific ways to improve schools that were in "underachieving" status. One school conducted an analysis of state and district assessment results to identify students with particular weaknesses. School staff found that high percentages of fifth-grade students were performing below proficiency in vocabulary and reading comprehension. This knowledge shaped the school's intervention, which involved implementing a mandated corrective-reading effort during the school day, a guided reading program after school, and the use of a 3-hour support aid. The author attributed the increased academic success of students in this school to these interventions and data-based curriculum changes.

A 2010 study from the U.S. Department of Education found that the most common uses of data in 36 case-study schools were school improvement planning, curricular decisions, and grouping for instruction or intervention. Other instructional responses were also noted in the previous literature:

- Making pacing decisions (Kuhs, Porter, Floden, & Freeman, 1985)
- Prioritizing instructional time (Brunner et al., 2005)
- Reteaching or reviewing concepts for the entire class (Blanc et al., 2010)
- Regrouping students (Blanc et al., 2010)
- Targeting students for intervention within the classroom (Henke, 2005; Snipes et al., 2002)
- Targeting students for intervention outside the classroom (Henke, 2005; Kuhs et al., 1985)

- Understanding the instructional effectiveness of individual lessons (Halverson, Prichett, & Watson, 2007; Supovitz & Klein, 2003)

- Informing classroom assignments (Snipes et al., 2002)

- Setting classroom, school, and district goals (Allinder & Oats, 1997; Kuhs et al., 1985; Marsh et al., 2006; Marshall, 2008; Niemi et al., 2007; Young, 2006)

For the purposes of our study, we defined the following key elements of classroom-level **Instructional Responses:**

- Establishing and/or adjusting student groupings

- Changing the scope and sequence

- Reviewing, reteaching, and/or altering the manner in which teachers introduce new topics

- Aligning lessons for reteaching
  - reflection on teaching methods (overall and for individual concepts)
  - assessment of teachers' own professional development needs
  - differentiated instruction

- Providing supplemental resources to "target" students
  - supplemental interventions and support for struggling students
  - additional attention to "bubble" students

Interrelationships Among the Key Dimensions of Data Use

The theory of action outlines a series of relations among the Key Dimensions. These relations, represented by arrows, capture how data use at the broad general levels (on the left side of the theory of action) relate to more individual or local data-use practices by schools, principals, and ultimately teachers (on the right side of the theory of action). To examine data use in urban schools and to test the nature of these relationships, we are conducting an empirical study in four urban districts.

Overview of the Study

The four districts participating in this study were selected on the basis of multiple inclusion criteria, drawing on data from a district-level survey that we administered to all 67 member districts of the Council of the Great City Schools. Specifically, we used the district survey data and additional supplemental information to identify districts that met four criteria: (1) the district had administered interim assessments continuously for the past 3 years; (2) the district planned to continue administering interim assessments for at least the next several school years; (3) the district administered interim assessments at least three times in a school year; and (4) the district data system had the capacity to meet the requirements of our quantitative study that will link school- and classroom-level data-use practices with student achievement. The selected districts also had to be willing to participate in the in-depth study, which included both a school-level survey component and a 2-day site visit.

To measure the Key Dimensions of data use in the theory of action, we administered surveys during the 2009–2010 school year to teachers in a sample of 183 randomly selected elementary and middle schools in the four districts. Teachers in the sampled schools were invited to participate in the surveys if they taught reading or mathematics in grades 4, 5, 7, or 8. Although not included in this report, all principals (and assistant principals, where appropriate) of participating schools were also asked to complete the surveys.

The overarching goal of the study is to address two primary questions regarding data-driven instruction: (1) What is the relationship between teachers' use of interim assessment data and their effectiveness at raising student achievement? (2) What is the relationship between school policies, practices, and resources for data-driven instruction and student achievement?

The links among the data use practices measured with teacher and principal surveys and student achievement will be the focus of future reports from this study.

In this report, we focus on the interrelationships among the data-use practices as they occur at different levels of the system, as reported by teachers, to provide a preliminary test of the theory of action. For this preliminary study, we posited three broad hypotheses about the relationships among data-use practices:

Hypothesis 1: *Context* (including assessment context, instructional context, and the district and school level data culture) will be positively related to (a) *Supports for Data Use* (as indicated by functional/easy-to-use data infrastructure, more organizational support, and high-quality professional development/staffing); (b) *Working With Data* (including collaboration and specific ways of making sense of student data); and (c) *Instructional Responses* (the extent to which teachers change their instruction in response to data).

Hypothesis 2: *Supports for Data Use* will be associated with (a) *Working With Data*, and (b) *Instructional Responses*.

Hypothesis 3: Finally, the extent to which teachers *Work With Data* will be positively related to *Instructional Responses*.

We also acknowledge the presence of potential relationships among the elements within Key Dimensions. For example, teacher collaboration may be positively related to teacher-principal or teacher-student collaboration. We therefore empirically tested the links within the Key Dimensions in the theory of action (context, supports for data use, working with data, and instructional responses) as well as the between-dimension relationships outlined in Hypotheses 1–3.

We tested these hypotheses and within-dimensions relationships empirically by using survey data from reading and mathematics teachers in grades 4 and 5 in the four participating districts.

Methods

Data Collection Procedures

Data collection occurred over the course of the 2009–2010 school year. We administered teacher surveys at three time intervals that were meant to coincide with the end of district interim assessments—the first in fall 2009, the second in winter 2010, and the third in spring 2010. Table 1 displays the administration dates of the first two waves of surveys in each district—these Wave 1 and Wave 2 teacher survey data are the focus of the analyses reporting in the sections that follow.[3]

Table 1. Survey Administration Dates by District

	District A	District B	District C	District D
Wave 1				
Date Survey Opened	11/16/2009	11/30/2009	11/3/2009	11/18/2009
Date Survey Closed	12/8/2009	12/18/2009	11/30/2009	12/9/2009
Wave 2				
Date Survey Opened	2/9/2010	3/1/2010	3/1/2010	3/17/2010
Date Survey Closed	3/2/2010	3/29/2010	3/29/2010	4/14/2010

All surveys were administered online. An initial email was sent to the district email address of every teacher in our sample, inviting them to participate in our study. Email reminders were sent weekly after the initial survey invitation to nonrespondents. Other follow-up strategies included postcard reminders approximately 1 week after the administration of the survey and follow-up calls made 1 week before the end of the survey. Upon completion of the online survey, teachers were sent a $25 Amazon.com® gift card. In the second and third waves of administration, survey respondents were entered into a raffle for an additional $100 Amazon.com® gift card. Five teachers were selected at random from each district to win, with the exception of District A, where three teachers were selected because of its smaller sample size.

Sample

Among the fourth- and fifth-grade teachers invited to participate in the study, 71 percent responded in Wave 1, 70 percent responded in Wave 2, and 60 percent responded in both waves (see Table 2).

Table 2. Sample Size and Response Rates by Wave

	Wave 1			Wave 2			Both Wave 1 and Wave 2		
	Invited	Responded		Invited	Responded		Invited	Responded	
Teachers	N	N	%	N	N	%	N	N	%
All 4th and 5th	810	574	71%	812	566	70%	809	483	60%

Because elementary school teachers often teach both mathematics and reading, we invited all fourth- and fifth-grade teachers in the participating schools to take the survey and respond about both subjects. Most of the teachers who responded did answer items about both mathematics and reading.

[3] The current report presents findings on Wave 1 and Wave 2 of the grades 4 and 5 teacher surveys only. Analyses have not yet been conducted that incorporate Wave 3 data and the teachers of grades 7 and 8, but they will be included in future reports.

For Wave 1, of the 574 teachers who responded, 495 taught mathematics and were included in the mathematics teacher sample, and 535 taught reading and were included in the reading teacher sample. A total of 473 of the 574 Wave 1 respondents taught both mathematics and reading; therefore, 82 percent of the sample was included in both the mathematics and reading analyses.

For Wave 2, of the 566 teachers who responded, 476 were included in the mathematics teacher sample, 523 were included in the reading teacher sample, and 449 (79 percent) were included in both samples.

Finally, out of the 483 teachers that responded in both Waves 1 and 2, 415 were included in the mathematics teacher sample, 453 were included in the reading teacher sample, and 396 (82 percent) were included in both samples.

In the next section and in Table 3, we describe the demographic characteristics of the teachers who are included in the survey sample used for this preliminary analysis.

Mathematics Teachers

Of the fourth- and fifth-grade teachers who taught mathematics and responded to the surveys in Wave 1, Wave 2, or in both waves, 85 percent were female. The mathematics teacher sample was 77 percent white, 18 percent African American, and 13 percent Hispanic.

Approximately 36 percent had completed their bachelor’s degree. Forty percent had also completed a master’s degree. An additional 25 percent had completed some type of post-master’s work. On average, they had about 12 years of teaching experience overall, with 6 in their current school.

Reading Teachers

Of the fourth- and fifth-grade teachers who taught reading and responded to the surveys (in Wave 1, Wave 2, or both), 85 percent were female. Seventy-eight percent were white, 18 percent were African American, and 13 percent of respondents reported being of Hispanic ethnicity.

With respect to educational achievement, 36 percent of respondents held a bachelor’s degree and 40 percent had also completed their master’s degree. Another 24 percent of respondents had completed some sort of post-master’s work (e.g., professional diploma, certificate of advanced graduate studies). On average, respondents had been teaching for approximately 12 years, with 9 years of experience in their current school.

Table 3. Demographic Characteristics of Fourth- and Fifth-Grade Teacher Sample

	Mathematics		Reading	
	Number	**Percent**	**Number**	**Percent**
Highest degree obtained				
Below bachelor's degree	0	0%	1	0%
Bachelor's degree	178	36%	198	36%
Master's degree	201	40%	217	40%
Educational Specialist or Professional Diploma	54	11%	60	11%
Certificate of Advanced Graduate Studies	60	12%	68	12%
Doctorate or Professional Degree (Ph.D., Ed.D.)	8	2%	8	1%
Gender (# Female)	425	85%	466	77%
Race				
White	346	77%	381	78%
Black	81	18%	86	18%
Asian	4	1%	4	1%
American Indian	1	0%	1	0%
Native Hawaiian or Pacific Islander	4	1%	4	1%
Multiracial	11	3%	11	2%
Ethnicity				
Latino/Latina	65	13%	68	13%
Missing race/ethnicity information	109	20%	118	20%
	Mean	**SD**	**Mean**	**SD**
Years of teaching experience (total)	11.69	8.80	11.70	6.10
Years of teaching experience (in current school)	6.00	5.36	8.89	5.39

Sample Size - Mathematics: N = 556 teachers; Reading: N = 605 teachers

Measures of Data Use Practices

Waves 1 and 2 Teacher Surveys

We developed teacher surveys to measure specific practices within the four hypothesized Key Dimensions of data use: (1) Context, (2) Supports for Data Use, (3) Working With Data, and (4) Instructional Responses. To create measures of the Key Dimensions, the study team drew from previously used survey instruments to use items with known psychometric properties. Where necessary, we developed new items to measure concepts within the theory of action that were not found in existing instruments. In summer 2009, we used an in-depth cognitive laboratory process to pilot the items and correct any issues with item sets before administering the first wave of surveys in fall 2009. The surveys were administered three times to the same teachers; each survey was designed to take approximately 20–30 minutes to complete.

Although we administered multiple surveys to the same respondents, the teacher surveys were not identical in each wave. Most items were repeated in the Wave 1 and Wave 2 surveys. However, some item sets were included only in either Wave 1 or Wave 2. For example, questions about data coaches were included only in Wave 2. The differences between Wave 1 and Wave 2 survey content were made to reduce survey length and burden on teachers. Table A-1 provides the timeline of when we measured each construct in the theory of action.

In Wave 1 of the survey, we asked elementary school–level teachers about how they used data *across* reading and mathematics (knowing that most, but not all, teach both subjects). However, in conversations with teachers of grades 4 and 5 during site visits, it became clear to the study team that in the participating districts, the mathematics and reading assessment process is different, so teachers' perceptions of Key Dimensions such as alignment and supports may be very different for the two subjects. Therefore, in Wave 2 of the survey, some items were restructured so that teachers could provide separate responses for reading and mathematics. The content that is separate for reading and mathematics in Wave 2 includes items about assessment context, data infrastructure, collaboration with parents around data, attention to data, making sense of data, and instructional responses.

Creating Scale Scores

Before any analyses were conducted, we first created scale scores to combine information from multiple item sets from the teacher surveys. We created scale scores separately for Wave 1 and Wave 2. For example, for Assessment Context, every teacher has both a Wave 1 Assessment Context Scale Score and a Wave 2 Assessment Context Scale Score. Because some items were administered only in either Wave 1 or Wave 2, some scale scores were created only for the available wave (see Table A-1 in Appendix A). We constructed the scale scores by averaging multiple items that measured the same construct in the theory of action. For example, if a construct was measured by 10 Likert-type items, the average of each teacher's 10 responses was used as his or her scale score.[4] Table 4 provides information about when each scale score was measured, by wave. More detailed information about specific items in each scale is shown in Appendix A.

Table 4. Scale Scores by Survey Administration Wave

Scale Scores	Timeline	
	Wave 1	Wave 2
Assessment Context	X	X
Instructional Context	X	Not Administered
District Data Culture	X	Not Administered
School Data Culture	x	X
Data Infrastructure	X	X
Professional Development	X	X
Data Coaches	Not Administered	X
Organizational Support	X	X
Collaboration With Coach	X	X
Collaboration With Teachers	X	X
Collaboration With Principals	X	X
Collaboration With Parents	X	X
Collaboration With Students	X	X
Attention to Data	X	X
Making Sense of Data	X	X
Instructional Responses	Not Included in Analyses	X

Note. For the dependent variable (instructional responses), only Wave 2 data were used to maintain temporal precedence of predictors in the model.

[4] Average scale scores were created for teachers if they answered at least two items included in a scale, therefore allowing respondents with missing data at the item level to still have scale scores.

Analytic Approach

To empirically test the theory of action, we used structural equation modeling (SEM). In this statistical approach, latent variables are created that combine information from multiple surveys or scales of surveys. Observed variables are directly measured (such as with a survey, observation, or interview). Observed variables are directly measured (such as with a survey, observation, or interview). Latent variables, in contrast, represent underlying constructs that are measured using multiple observed variables. For example, a latent variable of socio-economic status may be made up of the observed variables of education, income, and professional status.

Our analysis proceeded in a two-step fashion common when using SEM. First, combinations of observed variables (i.e., the scale scores from the teacher surveys) were used as indicators of latent variables; second, the proposed relationships between latent variables were tested to see whether these latent variables are related to each other in ways that are statistically significant. We estimated separate models for mathematics and reading.

Results

We examined the descriptive statistics for each scale score prior to conducting the SEM analysis. The dependent variable, teacher's *Instructional Responses*, ranged from 0 to 4, with an average score of 1.62 for both mathematics and reading teachers. On the response scale for these items, this mean relates to minor instructional change. Thus we see that on average, teachers reported low levels of instructional responses to interim assessment data. In contrast, teachers reported relatively positive perceptions of the data infrastructure in their districts. On the Data Infrastructure scale score, the range was 0 to 4.25 and the mean was 3.10. This average score translates to general agreement that the district and school data infrastructures are easy to use. Descriptive statistics for all scale scores used in the mathematics and reading analysis models are shown in Table A-2 in Appendix A.

Measurement Models: Creating Latent Variables

Our first step was to use the scale scores from the teacher surveys to build latent variables to measure each of the four Key Dimensions of data use and attempted to create a model that would only use one latent variable for each of the four Key Dimensions (Context, Supports for Data Use, Working With Data, and Instructional Responses). However, this model did not fit the data, meaning that it was not the best way to test how the key dimensions of data use were related to each other.

To be able to see how different types of data use were related to each other, our next step was to model each of the four key dimensions with their sub-components. For example, according to the theory of action; Data Infrastructure, Staffing, and Organizational Support are all elements of Supports for Data Use. Instead of assuming all these Supports for Data Use (Data Infrastructure, Staffing, Organizational Support) are related in the same way, we modeled all three elements separately. By splitting up the Key Dimensions in this way, the model better fit the data, suggesting that the relationships between different aspects of data use could be estimated with more confidence.

Therefore our final SEM included seven observed and latent variables to understand how different aspects of data use were related to each other:

- 6 latent variables:
    - *Context* was represented by a single latent variable, Assessment Context;
    - *Supports for Data Use* was represented by three latent variables (Data Infrastructure, Staffing, Organizational Support);
    - *Working With Data* was represented by two latent variables (Collaboration and Attention to Data).

- 1 observed variable of teachers' instructional response to data
    - There was no latent variable of *Instructional Responses*, but only the observed Wave 2 scale score was used in analyses.

Table 5 lists the different scale scores as well as the specific survey wave and item they came from, which make up each latent variable in the measurement model. Table A-3 in Appendix A gives more information about the latent variables, including factor loadings.

Table 5. Scale Score Indicators of Each Latent Variable

Key Dimension	Latent Variable	Scale Score Indicator
Context	Context	Assessment Context Wave 1
		Assessment Context Wave 2
		Instructional Context
		District Data Culture Wave 1
		School Data Culture Wave 1
		School Data Culture Wave 2
Supports for Data Use	Data Infrastructure	Data Infrastructure Wave 1
		Data Infrastructure Wave 2
	Staffing	Professional Development Wave 1
		Professional Development Wave 2
		Data Coach Wave 2
	Organizational Support	Organizational Supports - Grade Teams Wave 1
		Organizational Supports - Grade Teams Wave 2
Working with Data	Collaboration	Collaboration With Coach Wave 1
		Collaboration With Coach Wave 2
		Collaboration With Teachers Wave 1
		Collaboration With Teachers Wave 2
		Collaboration With Principal Wave 1
		Collaboration With Principal Wave 2
		Collaboration With Parents Wave 1
		Collaboration With Parents Wave 2
		Collaboration With Students Wave 1
		Collaboration With Students Wave 2
	Attention to Data	Attention to Data - Wave 1
		Attention to Data - Wave 2
		Making Sense of Data - Wave 1
		Making Sense of Data - Wave 2
Instructional Responses	N/A	Instructional Responses Wave 2

Note. Instructional Responses is an observed, not latent, variable.

Path Models

As the second step in our analytic approach, we used SEM to test the hypothesized relationships among the Key Dimensions in our theory of action. Path analysis models using SEM test the existence of theoretical relationships by using multiple regression between both latent and observed variables. We constructed a path model (one for mathematics and one for reading) that tests the links among the six latent variables and the relationships between each of them and teacher instructional responses.[5]

We estimated the paths from left to right as depicted in the theory of action presented in Exhibit 1, such that (1) *Context* was potentially related to all other latent variables and the outcome; (2) the *Supports for Data Use* variables were potentially related to the *Working With Data* variables and the outcome; and (3) the outcome (*Instructional Response*) was potentially predicted by all six latent variables.

Exhibits 2 and 3 depict the path models for mathematics and reading, respectively. These figures show only the paths that were statistically significant. (Full results, including all paths and covariances for each model, are presented in Tables B-1–B-3 in Appendix B).[6]

Mathematics

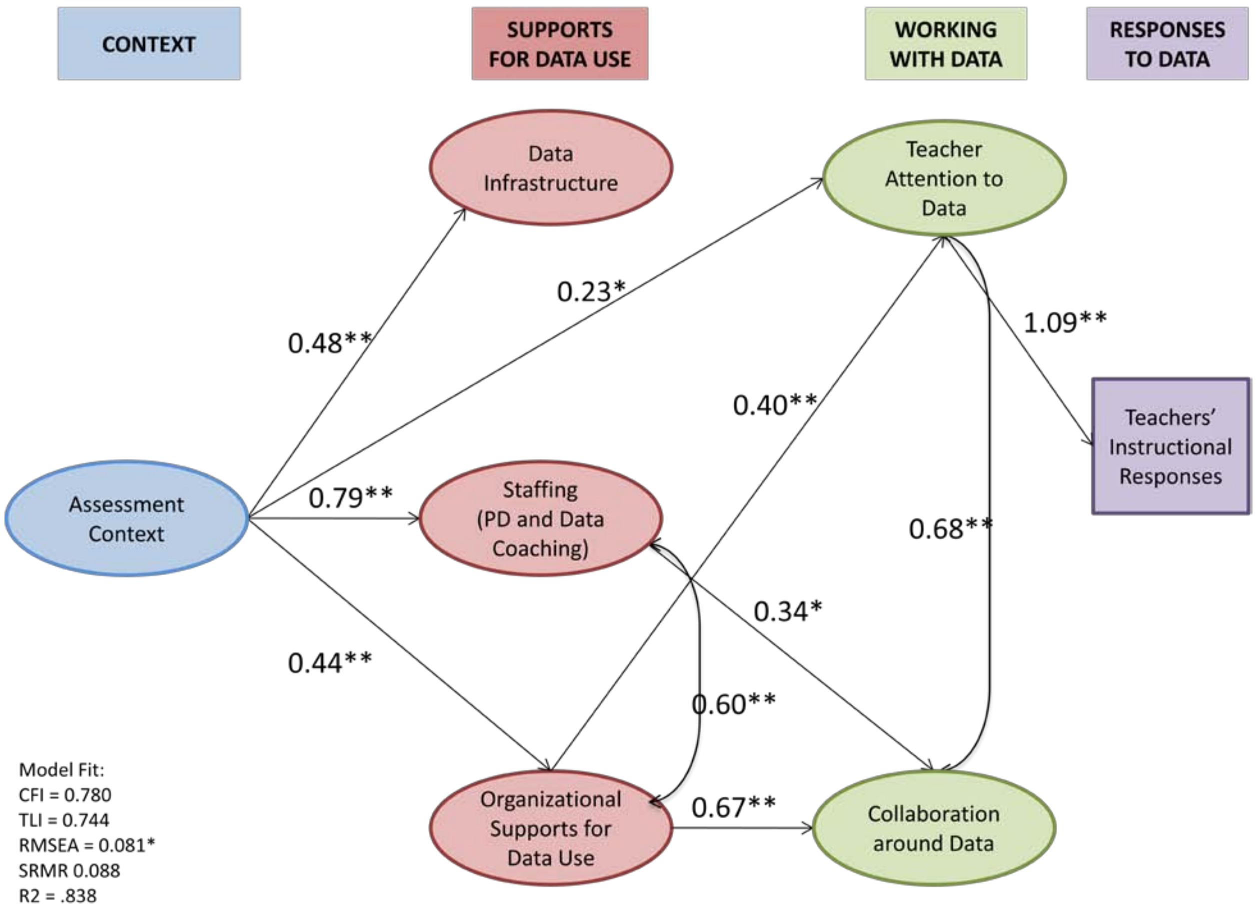
Results based on teacher surveys about mathematics interim assessments are presented in Figure 2. In general, the model supports the theory of action. Statistically significant positive relationships exist between latent variables that represent the Key Dimensions as hypothesized and indicate that the constructs in the theory of action are inter-related, as expected.

---

[5] The path analytic models presented in this study are correlational and reveal relationships or associations. No control group was included in the study; therefore, the analysis cannot support causal conclusions. The fact that our measurement of the theory of action constructs depends entirely on subjective self-report measures further limits our ability to make causal statements about whether and how the Key Dimensions of data use influence teachers' instructional responses to data.

[6] Model fit for both measurement and path models in reading and mathematics were highly similar and indicated moderate model fit. CFI and TLI ranged between 0.7 and 0.8, which are below recommended values of 0.95. RMSEA of 0.08 is borderline, and SRMR of 0.09 is above recommended values of 0.05. However, the statistically significant R^2 value of 0.83 for both the reading and mathematics analyses indicates that a substantial percentage of variance in the outcome is accounted for by the models (83 percent).

Exhibit 2. Path Analytic Model for Fourth- and Fifth-Grade Mathematics Teachers



Note. * *p* < 0.05, ** *p* < 0.01

Statistically significant paths from assessment context (including data culture to all three types of supports for data use—data infrastructure, staffing, and organizational supports for data use) suggest that *Context* is related to *Supports for Data Use* in ways entirely consistent with our theory of action.

Interestingly, and contrary to our hypotheses, the three types of supports for data use were expected to be significantly interrelated and they are not; staffing and (concrete) organizational supports for data are related, but neither is related to data infrastructure. This finding suggests that data infrastructure may be independent from other facets of supports for data use provided by districts. Collaboration around data is significantly predicted by both staffing ($\beta = 0.34$, $p < .01$) and organizational supports for data use ($\beta = 0.67$, $p < .001$). In turn, collaboration is highly correlated with teacher attention to data ($r = 0.68$, $p < .001$), suggesting, as expected, that teachers who collaborate with others around data also spend more time reviewing and making sense of data.

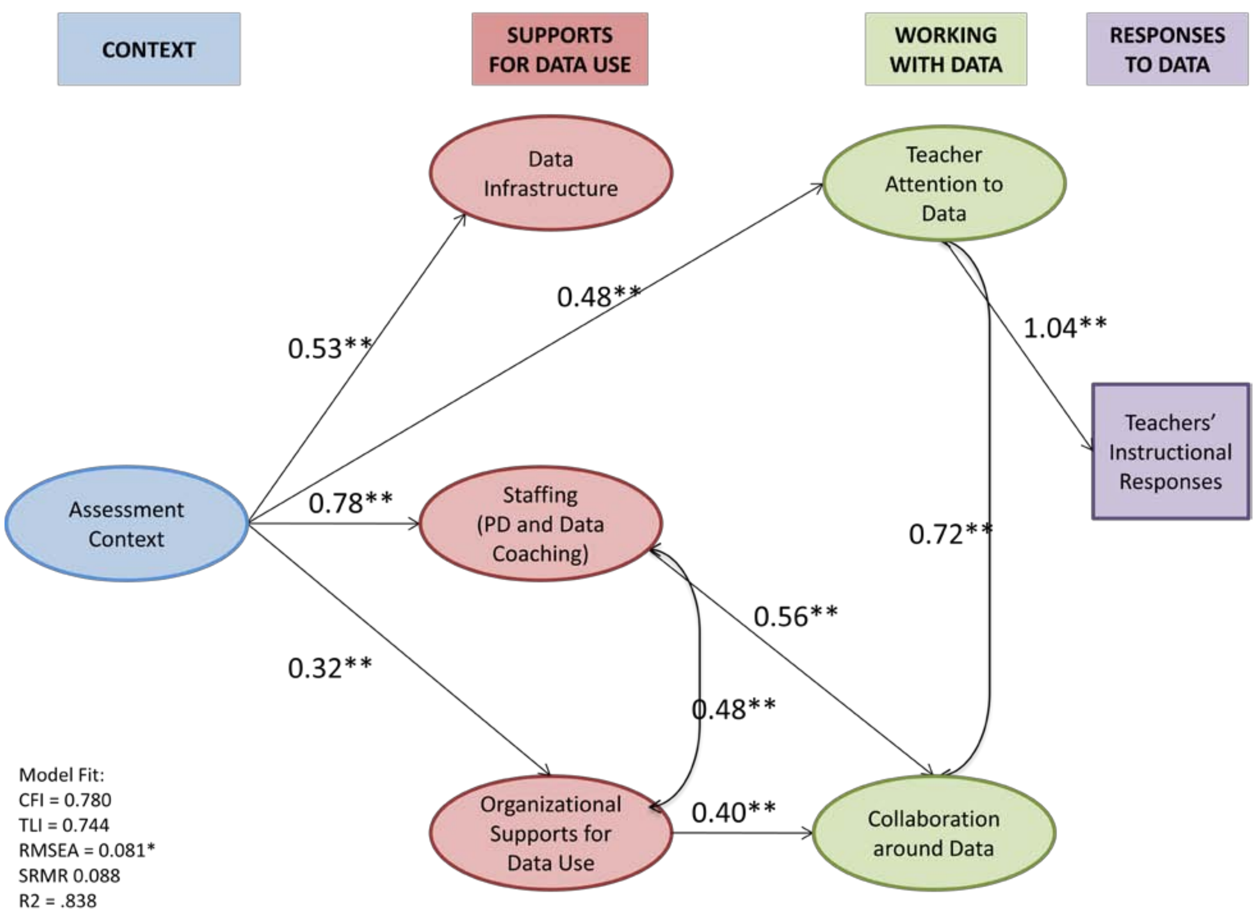
The interrelationships between teacher attention to data and the other variables indicate that, as expected, *Context* and *Supports for Data Use* (specifically, concrete organizational supports), are related to the degree to which teachers spend time reviewing and making sense of data. We also see that the relationship between *Context* and *Working With Data* is not entirely mediated through *Supports for Data Use*, but rather there is a direct effect of *Context* on Teacher Attention to Data.

The only direct predictor of teachers' *Instructional Responses* was Teacher Attention to Data ($\beta = 0.72$, $p < .001$). That is, of the six latent variables, only teachers' attention to data directly predicted the degree to which teachers report using data to change their instruction. Contrary to our hypotheses, Collaboration Around Data did not significantly predict teachers' instructional responses to data directly but rather may be indirectly related to *Instructional Responses* through a link with Teacher Attention to Data.

Reading

Results for our analysis of teacher survey responses about reading interim assessments are similar to those for mathematics and are presented in Exhibit 3.

Exhibit 3. Path Analytic Model for Fourth- and Fifth-Grade Reading Teachers



Note. ** $p < 0.01$

In reading as with mathematics, the results support the relationships proposed in the theory of action. Specifically, we found that *Context* is related to *Supports for Data Use*, *Supports for Data*

Use is related to higher levels of *Working With Data*, and within *Working With Data,* Teacher Attention to Data is the only significant predictor of teachers' *Instructional Responses* to data.

The only notable difference between the reading and mathematics models was in the relationship between Organizational Supports for Data Use and Teacher Attention to Data. Although this path was statistically significant in the mathematics model (β = 0.40, p < 0.001), it was not in the reading model (β = 0.08, p = 0.32). All other paths that were significant in the mathematics model were also significant in the reading model; however, the magnitude of the path coefficients varied.

Conclusions and Recommendations

The literature about using periodic assessment data to improve instruction is not clear cut. Prior evidence suggests that systematic measurement of student learning can improve achievement under certain conditions. There also is some consensus in the literature about the facilitating factors and barriers to data use. However, some studies suggest that interim assessments do not improve student achievement, and the literature does not provide clear or complete answers about what educators should specifically do with interim assessment data to increase student learning. As stated by Herman and colleagues (2008) in their study of data use in urban public schools, "The field of education has not necessarily developed a strong and concrete set of best practices around data use" (p. 40).

Using our theory of action as a framework, we see that the results of our study begin to point to at least some categories of practice that appear important for facilitating teachers' use of data to inform their instruction. Our results suggest that teachers in urban districts who perceive the presence of facilitating contextual conditions and concrete supports for data use are more likely to actually review interim assessment data and, in turn, to use those data to change their instruction.

Of note, we specifically found that teachers who engage in more collaboration around data engage in more practices related to reviewing data. Teacher attention to data, in turn, is a significant and positive predictor of their instructional responses. These findings suggest that teachers who spend more time reviewing data are more likely to adjust their instructional strategies and educational decision making in their classrooms.

In combination, the results indicate that districts interested in increasing the likelihood that teachers will use data to inform their instruction may find it useful to provide concrete supports, including structured time for teacher to collaborate around and review interim assessment data.

We also found that the assessment context at the district and school levels is an important predictor of the other Key Dimensions of interim assessment data use and thus should not be overlooked. Context included aspects of assessment quality and alignment, as well as a "culture of data use." When results were broken down by district, we noted that teachers' reported perceptions of the context varied by district, but in all four districts, teachers generally reported that the context represented supporting conditions for using data. It may be useful for districts implementing a new interim assessment program (or seeking to improve their existing program), to spend time strengthening their contextual conditions and fostering a culture of data use within and across the district and schools.

In this report, we focused on teacher's instructional responses as the outcome. In future reports we will examine whether data-based instructional responses are in turn related to improvements in student achievement on state assessments. Thus, future reports from this study will extend the current findings to examine the relationships between the Key Dimensions of data use and student achievement, as measured by state assessments in both mathematics and reading.

References

Abelman, C., Elmore, R. F., Even, J., Kenyon, S., & Marshall, J. (1999). *When accountability knocks, will anyone answer?* Philadelphia: Consortium for Policy Research in Education and the University of Pennsylvania Graduate School of Education.

Allinder, R. M., & Oats, R. G. (1997). Effects of acceptability on teachers' implementation of curriculum-based measurement and student achievement in mathematics computation. *Remedial and Special Education*, *18*(2), 113–120.

American Recovery and Reinvestment Act of 2009, Pub. L. No. 111-115, § 3, 123 Stat. 182.

Anderson, S., 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.

Black, P. J., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2002). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, *86*(1), 8–21.

Black, P. J., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy, and Practice*, *5*(1), 7–75.

Black, P. J., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–144, 146–148.

Blanc, S., Christman, J. B., Liu, R., Mitchell, C, Travers, E., & Bulkley, K. E. (2010). Learning to learn from data: Benchmarks and instructional communities. *Peabody Journal of Education*, *85*(2), 205–225.

Brookhart, S. M. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education*, *8*(2), 153–169.

Brunner, C., Fasca, C., Heinze, J., Honey, M., Light, D., Mandinach, E., et al. (2005). Linking data and learning: The Grow Network study. *Journal of Education for Students Placed at Risk*, *10*(3), 241–267.

Bulkley, K. E., Christman, J. B., Goertz, M. E., & Lawrence, N. R. (2010). Building with benchmarks: The role of the district in Philadelphia's benchmark assessment system. *Peabody Journal of Education*, *85*(2), 186–204.

Clune, W. H., & White, P. A. (2008). Policy effectiveness of interim assessments in Providence *Public Schools* (WCER Working Paper No. 2008-10). Madison, WI: Wisconsin Center for Center for Education Research. Retrieved December 16, 2008, from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/41/b4/25.pdf

Data Quality Campaign (2009). *The next step: Using longitudinal data systems to improve student success.* Washington DC: Author. Retrieved October 9, 2009, from http://www.dataqualitycampaign.org/about

Datnow, A., Park, V., & Kennedy, B. (2008). *Acting on data: How urban high schools use data to improve instruction*. Los Angeles: University of Southern California, Center of Educational Governance.

Datnow, A., Park, V., & Wohlsetter, P. (2007). *Achieving with data: How high-performing school systems use data to improve instruction for elementary students*. Los Angeles: University of Southern California, Center on Educational Governance.

David, J. L. (2008). Pacing guides. *Educational Leadership, 66*(2), 87–88.

Davis, L. B., & Fuchs, L. S. (1995). Will CBM help me learn? Students' perception of the benefits of curriculum-based measurement. *Education & Treatment of Children*, *18*(1), 19–32.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*, 219–232.

Deno, S. L., & Fuchs, L. S. (1987). Developing curriculum-based measurement systems for data-based special education problem solving. *Focus on Exceptional Children*, *19*, 1–16.

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.

Fuchs, L. S., Deno, S. L., & Mirkin, P. K. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement, and student awareness of learning. *American Educational Research Journal*, *21*(2), 449–460.

Fuchs, L. S., & Fuchs, D. (1990). The role of skills analysis in curriculum-based measurement in math. *School Psychology Review*, *19*(1), 6–22.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (2007). Using curriculum-based measurement to inform reading instruction. *Reading and Writing*, *20*(6), 553–567.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal*, *28*(3), 617–641.

Halverson, R., Prichett, R. B., & Watson, J. G. (2007). *Formative feedback systems and the new instructional leadership*. Madison: University of Wisconsin.

Hamilton, L. (2005). *Assessment as a policy tool.* Santa Monica, CA: RAND Corporation.

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: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/.

Hayward, L., Priestley, M., & Young, M. (2004). Ruffling the calm of the ocean floor: Merging practice, policy and research in assessment in Scotland. *Oxford Review of Education*, *30*(3), 397–415.

Henderson, S., Petrosino, A., Guckenburg, S., & Hamilton, S. (2007). *Measuring how benchmark assessments affect student achievement* (Issues & Answers Report, REL 2007–No. 039). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved June 3, 2009, from http://ies.ed.gov/ncee/edlabs/projects/project.asp?ProjectID=43

Henderson, S., Petrosino, A., Guckenburg, S., & Hamilton, S. (2009). *A second follow-up year for measuring how benchmark assessments affect student achievement* (REL Technical Brief, REL Northeast and Islands 2007–No. 002). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved June 3, 2009, from http://ies.ed.gov/ncee/edlabs/projects/project.asp?ProjectID=43

Henke, K. G. (2005). From vision to action: How school districts use data to improve performance. Washington, DC: Consortium for School Networking.

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, *89*(2), 140–145.

Heritage, M., & Bailey, A. L. (2006). Assessing to teach: An introduction. *Educational Assessment, 11*(3/4), 145 –148.

Herman, J., & Gribbons, B. (2001). *Lessons learned in using data to support school inquiry and continuous improvement: Final report to the Stuart Foundation.* Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Herman, J. L., Yamashiro, K., Lefkowitz, S., & Trusela, L. A. (2008). Exploring data use and school performance in an urban public school district: Evaluation of Seattle Public Schools' comprehensive value-added system (CRESST Report 742). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved June 3, 2009, from http://www.eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?_npb=true&_&ERICExtSearch_SearchValue_0=ED503302&ERICExtSearch_SearchTyp0=no&accno=ED503302

Hintze, J. M., Daly III, E. J., & Shapiro, E. S. (1998). An investigation of the effects of passage difficulty level on outcomes of oral reading fluency progress monitoring. *School Psychology Review*, *27*(3), 433–447.

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.

Kerr, K. A., Marsh, J. A., Ikemoto, G. S., Darilek, H., & Barney, H. (2006). Strategies to promote data use for instructional improvement: Actions, outcomes, and lessons from three urban districts. *American Journal of Education*, *112*(4), 496–520.

Kuhs, T., Porter, A., Floden, R., & Freeman, D. J. (1985). Differences among teachers in their use of curriculum-embedded tests. *The Elementary School Journal*, *86*(2), 141–153.

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.

Long, L., Rivas, L. M., Light, D., & Mandinach, E. B. (2008). The evolution of a homegrown data warehouse: TUSDStats. In E. B. Mandinach & M. Honey (Eds.), *Data-driven school improvement: Linking data and learning* (pp. 209–232). New York: Teachers College Press.

Love, N. (2004). Taking data to new depths. *Journal of Staff Development*, *25*(4), 22–26.

Marsh, J. A., Pane, J. F., & Hamilton, L. S. (2006). *Making sense of data-driven decision making in education: Evidence from recent RAND research.* Santa Monica, CA: RAND Corporation.

Marsh, J. A., McCombs, J. S., & Martorell, F. (2009, August). How instructional coaches support data-driven decision making: Policy implementation and effects in Florida middle schools. *Educational Policy* online. Retrieved September 30, 2009, from http://epx.sagepub.com/content/early/2009/08/03/0895904809341467.

Marshall, K. (2006). *Interim assessments: Keys to successful implementation.* New York: New Leaders for New Schools. Retrieved from http://www.marshallmemo.com/articles/Interim%20Assmt%20Report%20Apr.%2012,%2006.pdf

Marshall, K. (2008). Interim assessments: A user's guide. *Phi Delta Kappan*, *90*(1), 64–68.

Marston, D., Pickart, M., Reschly, A., Heistad, D., Muyskens, P., & Tindal, G. (2007). Early literacy measures for improving student reading achievement: Translating research into practice. *Exceptionality*, *15*(2), 97–117.

Mason, S. (2002). *Turning data into knowledge: Lessons from six Milwaukee Public Schools.* Madison: University of Wisconsin-Madison, Wisconsin Center for Education Research.

May, H., & Robinson, M. A. (2007). *A randomized evaluation of Ohio's Personalized Assessment Reporting System (PARS).* Philadelphia, PA: Consortium for Policy Research in Education.

Militello, M., Schweid, J., & Sireci, S. G. (2010). Formative assessment systems: Evaluating the fit between school districts' needs and assessment systems' characteristics. *Educational Assessment, Evaluation, and Accountability*, *22* (1), 29–52.

Murnane, R. J., Sharkey, N. S., & Boudett, K. P. (2005). Using student-assessment results to improve instruction: lessons from a workshop. *Journal of Education for Students Placed at Risk*, *10*(3), 269–280.

Nabors Ol´ah, L., Lawrence, N. R., & Riggan, M. (2010). Learning to learn from benchmark assessment data: How teachers analyze results. *Peabody Journal of Education*, *85*(2), 226–245.

Niemi, D., Vallone, J., Wang, J., & Griffin, N. (2007). *Recommendations for building a valid benchmark assessment system: Interim report to the Jackson Public Schools* (CRESST Report 723). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Retrieved December 16, 2008, from http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/41/d/c7.pdf

Perie, M., Marion, S., Gong, B., & Wurtzel, J. (2007, November). *The role of interim assessments in a comprehensive assessment system*. Washington, DC: The Aspen Institute.

Phillips, N. B., Hamlett, C. L., Fuchs, L. S., & Fuchs, D. (1993). Combining classwide curriculum-based measurement and peer tutoring to help general educators provide adaptive education. *Learning Disabilities Research & Practice, 8*(3), 148–156.

Quint, J. C., Sepanik, S., & Smith, J. K. (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: MDRC. Retrieved January 15, 2008, from http://www.mdrc.org/publications/508/full.pdf

Research for Action. (2009). *Making the most of interim assessment data: Lessons from Philadelphia*. Philadelphia: Author. Retrieved June 30, 2009, from http://pdf.researchforaction.org/rfapdf/publication/pdf_file/559/Christman_J_Making_the_Most_of_Interim_Assessment_Data_ES.pdf

Sharkey, N. S., & Murnane, R. J. (2006). Tough choices in designing a formative assessment system. *American Journal of Education*, *112*(4), 572–588.

Shepard, L. A. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, *63*(3), 66–71.

Snipes, J., Doolittle, F., & Herlihy, C. (2002). *Foundations for success: Case studies of how urban schools improve student achievement.* New York: MDRC.

Stecker, P. M., & Fuchs, L. S. (2000). Effecting superior achievement using curriculum-based measurement: the importance of individual progress monitoring. *Learning Disabilities Research and Practice*, *15*(3), 128–134.

Supovitz, J. A., & Klein, V. (2003). *Mapping a course for improved student learning: How innovative schools systematically use student performance data to guide improvement.* Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development. (2009). *Implementing data-informed decision making in schools: Teacher access, supports and use.* Washington, DC: Author.

U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. (2010). *Use of education data at the local level from accountability to instructional improvement.* Washington, DC: Author.

Wayman, J. C., & Stringfield, S. (2006). Technology-supported involvement of entire faculties in examination of student data for instructional improvement. *American Journal of Education*, *112*(4), 549–571.

Wayman, J.C., Cho, V., & Johnston, M.T. (2007). *The data-informed district: A district-wide evaluation of data use in the Natrona County School District.* Austin, TX: The University of Texas.

Young, V. M. (2006). Teachers' use of data: Loose coupling, agenda setting, and team norms. *American Journal of Education*, *112*(4), 521–548.

APPENDIX A – MEASURES

Appendix A provides information about the measures used to capture information about interim assessment data use.

- Table A-1 presents the number of items that contributed to each scale score, along with examples of items in each scale score.
- Table A-2 presents descriptive statistics for each scale score.
- Table A-3 provides results for the measurement model used to create the latent variables.

Table A-1 presents specific information regarding how the scale scores were created, including number of items per wave, example items from the surveys, and response scales. “Not administered” indicates that the particular survey wave did not include any items used to measure the scale score.

Table A-1. Measuring the Theory of Action: Specific Items Included in Each Scale Score

Scale Scores	# Items		Example Items	Response Scale
	Wave 1	Wave 2		
Assessment Context	10	10*	*“The interim assessment is appropriately challenging for my students.” “The interim assessment is well-aligned with state and district standards.” “The interim assessment is well-aligned with state and district standards.”*	1= Strongly Disagree to 4 = Strongly Agree
Instructional Context	9	Not Administered	*How much do the following factors influence your teaching: “District's curriculum framework, standards, or guidelines,” “End-of-year state assessment scores”*	1= No Influence to 4= Major Influence
District Data Culture	4	Not Administered	*“The district sets clear, consistent goals for schools to use data for school improvement.” “The district has designated adequate resources (e.g. time, staff, money) to facilitate teachers' use of data.”*	1= Strongly Disagree to 4 = Strongly Agree
School Data Culture	7	7	*“Teachers in this school are continually learning and seeking new ideas.” “The principal at my school encourages teachers to make decisions based on data.”*	1= Strongly Disagree to 4 = Strongly Agree
Data Infrastructure	7	7*	*“Interim assessment data are easy to use.” “My school’s internet connection enables teachers to access the district interim assessment system online.” “Interim assessment results are reported to me in a timely manner.”*	1= Strongly Disagree to 4 = Strongly Agree
Professional Development	8	8	*“Over the last 12 months, how much did the professional development provided by your district’s central office emphasize the following? a) Linking student achievement data to classroom practice b) Incorporating interim assessment data into lesson planning.”*	1 = No emphasis to 4 = Major emphasis
Data Coaches	Not Administered	10	*“How much assistance does the data coach provide to your school in a) Organizing existing data, including interim assessment data b) Helping staff develop their own capacity to analyze data.”*	1 = No assistance; 2 = Limited assistance; 3 = Moderate assistance; 4 = Substantial assistance
Organizational Support	7	7	*How often does your school provide scheduled meeting time for teachers to conduct the following activities: “Review interim assessment results as grade-level teams,” “Meet with a data coach,” and “Discuss and share instructional strategies.”*	3=About once a week, 2 = 1-2 times per month, 1 = 1-2 times per quarter, 0 = My school does not provide time for this

Scale Scores	# Items		Example Items	Response Scale
	Wave 1	Wave 2		
Collaboration With Coach	4	4	*"How frequently do you review student interim assessment data with a content-area coach (e.g. math or reading coach)?"*	1 = Never to 4 = 1 or 2 times a week
Collaboration With Teachers	5	3	*"How frequently do you review student interim assessment data with classroom teachers?"*	1 = Never to 4 = 1 or 2 times a week
Collaboration With Principals	3	3	*"How frequently do you review student interim assessment data with school administrators?"*	1 = Never to 4 = 1 or 2 times a week
Collaboration With Parents	4	3*	*"How frequently do you review student interim assessment data with parents"*	1 = Never to 4 = 1 or 2 times a week
Collaboration With Students	5	3	*"On average how often do you use your interim assessment results to inform students of their progress?"*	0 = Never to 5 = Daily
Attention to Data	11	11*	*"To what extent do you use the following district interim assessment results? a) Percent of students scoring at or above the proficient level, b) Results for your class(es) c) Results on specific topics or skills (e.g., computation, applications, word recognition, grammar, etc.)"*	0 = Not made available in this way to 3 = Used extensively
Making Sense of Data	4	4*	*"How much have you used the latest interim assessment results to: a) Identify individual students who need remedial assistance, b) Diagnose learning problems."*	1 = Did not use in this way to 4 = Used extensively
Instructional Responses	Not Included in Analysis	23*	*"In the past 6 weeks, to what extent did you do the following to address the needs of students as a direct result of students' interim assessment scores in reading and math? a) Reviewed key concepts for the entire class b) Used same-level achievement groupings c) Provided individual assistance outside of class to address the needs of struggling students."*	1 = Did not use in this way to 4 = Used extensively

Note. * indicates separate items for mathematics and reading in Wave 2.

Table A-2 shows the range, mean, and standard deviation of each scale score created. Results are presented for the sample of mathematics teachers separate from reading teachers.

Table A-2. Descriptive Statistics for Each Scale Scores for the Mathematics and Reading Models

	Mathematics				Reading			
Scale Score	Min	Max	Mean	SD	Min	Max	Mean	SD
Assessment Context Wave 1	1.00	4.00	2.68	0.56	1.00	4.00	2.66	0.55
Assessment Context Wave 2	1.20	4.00	2.74	0.48	1.11	4.00	2.64	0.51
Instructional Context Wave 1	0.00	4.00	3.41	0.42	0.44	4.00	3.41	0.39
District Data Culture Wave 1	1.00	4.00	2.45	0.63	1.00	4.00	2.45	0.62
School Data Culture Wave 1	1.25	4.00	2.96	0.48	1.25	4.00	2.95	0.47
School Data Culture Wave 2	1.00	4.00	2.91	0.48	1.00	4.00	2.91	0.48
Data Infrastructure Wave 1	1.00	4.29	2.96	0.57	1.00	4.29	2.95	0.58
Data Infrastructure Wave 2	1.00	4.25	3.10	0.57	1.00	4.25	3.10	0.57
Data Coach Wave 2	1.10	4.70	3.06	0.81	1.10	4.70	3.08	0.80
Professional Development Wave 1	1.00	4.00	2.49	0.78	1.00	4.00	2.46	0.79
Professional Development Wave 2	1.00	5.00	2.50	1.27	1.00	5.00	2.53	1.27
Collaboration - Coach Wave 1	0.00	4.33	0.69	0.79	0.00	4.33	0.69	0.79
Collaboration - Coach Wave 2	0.00	4.25	1.21	0.89	0.00	4.25	1.22	0.89
Collaboration - Teacher Wave 1	0.00	3.44	1.58	0.67	0.00	3.44	1.57	0.67
Collaboration - Teacher Wave 2	0.00	4.50	1.81	0.79	0.00	4.40	1.88	0.78
Collaboration - Principal Wave 1	0.00	4.33	0.92	0.84	0.00	4.33	0.94	0.85
Collaboration - Principal Wave 2	0.00	4.67	1.26	0.83	0.00	4.67	1.29	0.84
Collaboration - Parents Wave 1	0.00	3.75	1.48	0.69	0.00	3.75	1.48	0.69
Collaboration - Parents Wave 2	0.00	4.00	1.63	0.71	0.00	4.00	1.62	0.70
Collaboration - Students Wave 1	0.00	4.60	1.63	1.06	0.00	4.60	1.61	1.05
Collaboration - Students Wave 2	0.00	4.50	2.13	1.07	0.00	4.50	2.11	1.05
Attention to Data - Wave 1	0.00	4.33	2.20	0.68	0.10	4.00	2.23	0.65
Attention to Data - Wave 2	0.33	3.36	2.10	0.62	0.00	3.50	1.92	0.62
Making Sense of Data - Wave 1	1.00	4.00	2.70	0.79	1.00	4.00	2.70	0.79
Making Sense of Data - Wave 2	1.00	4.00	2.52	0.78	1.00	4.00	2.50	0.78
Instructional Response - Total Wave 1	0.36	3.39	1.72	0.63	0.36	3.39	1.73	0.63
Instructional Response - Total Wave 2	0.36	4.00	1.62	0.65	0.28	4.00	1.62	0.64
Organizational Supports - Grade Teams Wave 1	0.00	3.00	1.62	0.69	0.00	3.00	1.53	0.67
Organizational Supports - Grade Teams Wave 2	0.00	3.00	1.60	0.76	0.00	3.00	1.60	0.76

Sample Size: Mathematics: N = 117 schools, 579 teachers; Reading: N = 117 schools, 629 teachers

Table A-3 provides information on the measurement models for both the mathematics and reading structural equation models. Factor loadings are provided for how each scale score loaded onto its respective latent variable. The first column, labeled "Latent Variable," indicates the latent variable that is described in that section of the table. The second column, labeled "Scale Score Variable," indicates the scale scores that were used to create each specific latent variable. The next three columns present the factor loadings, the standard error, and the associated *p*-value for each indicator of each latent variable. Standardized factor loadings can range in value from 0 to 1, where values closer to 1 indicate a stronger relationship between the variables. Acceptable factor loadings are values above 0.40. Factor loadings below 0.40 suggest poor fit for the latent variable. Factor loadings are considered statistically significant if the *p*-value is less than 0.05.

TableA-3. Measurement Model Factor Loadings for Mathematics and Reading

Latent Variable	Scale Score Variable	Mathematics Coefficient (*β*)	SE	*p*-value	Reading Coefficient (*β*)	SE	*p*-value
Context	Assessment Context Wave 1	0.604	0.06	0.000	0.584	0.06	0.000
	Assessment Context Wave 2	0.483	0.06	0.000	0.508	0.06	0.000
	Instructional Context	0.427	0.11	0.000	0.508	0.05	0.000
	District Data Culture Wave 1	0.723	0.04	0.000	0.678	0.05	0.000
	School Data Culture Wave 1	0.555	0.05	0.000	0.501	0.06	0.000
	School Data Culture Wave 2	0.474	0.07	0.000	0.452	0.07	0.000
Data Infrastructure	Data Infrastructure Wave 1	0.962	0.072	0.000	0.887	0.05	0.000
	Data Infrastructure Wave 2	0.667	0.064	0.000	0.715	0.06	0.000
Staffing	Professional Development Wave 1	0.687	0.04	0.000	0.663	0.04	0.000
	Professional Development Wave 2	0.561	0.05	0.000	0.587	0.05	0.000
	Data Coach Wave 2	0.434	0.06	0.000	0.462	0.06	0.000
Organizational Support	Organizational Supports - Grade Teams Wave 1	0.702	0.04	0.000	0.807	0.05	0.000
	Organizational Supports - Grade Teams Wave 2	0.801	0.03	0.000	0.724	0.04	0.000
Collaboration	Collaboration - Coach Wave 1	0.624	0.04	0.000	0.627	0.04	0.000
	Collaboration - Coach Wave 2	0.630	0.05	0.000	0.628	0.04	0.000
	Collaboration - Teacher Wave 1	0.717	0.04	0.000	0.716	0.04	0.000
	Collaboration - Teacher Wave 2	0.772	0.03	0.000	0.785	0.02	0.000
	Collaboration - Principal Wave 1	0.594	0.04	0.000	0.598	0.04	0.000
	Collaboration - Principal Wave 2	0.673	0.04	0.000	0.672	0.03	0.000
	Collaboration - Parents Wave 1	0.628	0.05	0.000	0.154	0.08	0.046
	Collaboration - Parents Wave 2	0.542	0.05	0.000	0.536	0.05	0.000
	Collaboration - Students Wave 1	0.694	0.04	0.000	0.651	0.04	0.000
	Collaboration - Students Wave 2	0.625	0.04	0.000	0.591	0.04	0.000
Working With Data	Attention to Data - Wave 1	0.782	0.04	0.000	0.580	0.06	0.000
	Attention to Data - Wave 2	0.685	0.06	0.000	0.761	0.04	0.000
	Making Sense of Data - Wave 1	0.730	0.03	0.000	0.633	0.05	0.000
	Making Sense of Data - Wave 2	0.645	0.07	0.000	0.754	0.05	0.000

Sample Size: Mathematics: N = 117 schools, 579 teachers; Reading: N = 117 schools, 629 teachers.

APPENDIX B – ESTIMATION METHODS AND HYPOTHESIS TESTING

Appendix B presents the results for the mathematics and reading structural equation models used to estimate the relationships between the different constructs in the theory of action.

- Table B-1 shows the results for the mathematics and reading structural equation path models.
- Table B-2 presents the covariances among latent variables in the mathematics structural equation model.
- Table B-3 presents the covariances among latent variables in the reading structural equation model.

Table B-1 presents the path coefficients for all paths tested in the structural equation models of interim assessment data use. The first column, labeled "Independent Variable," indicates the variable that is the predictor in the specific path/equation tested. The second column, labeled "Dependent Variable," indicates the variable that is the outcome in the specific path/equation tested. The next three columns, under the heading "Mathematics," present the path coefficients (β), the standard errors, and the associated p-values for tests of significance for the specific path/equation tested. Each coefficient (β) is the standardized estimate of the relationship between the dependent and independent variables. It represents the amount of change in the outcome variable in standard deviations units given a one standard deviation change in the predictor. Paths are considered statistically significant if the p-value is less than 0.05. For example, the path between Context and Data Infrastructure is significant, whereas the path between Data Infrastructure and Attention to Data is not significant. The last three columns present the findings for the reading model and include path coefficients (β), the standard error, and the associated p-value for the test of significance for each specific path/equation tested.

Table B-1. Results of Structural Equation Model of Paths Among Data Use Practices

Independent Variable	Dependent Variable	Mathematics Coefficient (β)	Mathematics SE	Mathematics p-value	Reading Coefficient (β)	Reading SE	Reading p-value
Context	Data Infrastructure	0.48	0.059	<0.001	0.53	0.06	<0.001
	Staffing	0.79	0.068	<0.001	0.78	0.068	<0.001
	Organizational Support	0.44	0.077	<0.001	0.32	0.08	<0.001
Context	Attention to Data	0.23	0.113	0.044	0.48	0.106	<0.001
Data Infrastructure		0.06	0.087	0.476	0.05	0.098	0.590
Organizational Supports		0.40	0.072	<0.001	0.06	0.063	0.360
Context	Collaboration	–0.12	0.151	0.416	–0.09	0.166	0.588
Data Infrastructure		0.07	0.065	0.296	0.56	0.167	0.001
Staffing		0.34	0.171	0.046	0.40	0.095	<0.001
Organizational Supports		0.67	0.098	<0.001	0.08	0.079	0.316
Context	Instructional Response	–0.19	0.184	0.305	–0.24	0.196	0.228
Data Infrastructure		–0.04	0.060	0.495	–0.02	0.065	0.707
Staffing		0.41	0.239	0.086	0.41	0.225	0.071
Organizational Supports		–0.05	0.149	0.730	–0.01	0.090	0.913
Attention to Data		1.09	0.132	<0.001	1.04	0.126	<0.001
Collaboration		–0.38	0.294	0.192	–0.28	0.215	0.186
Sample Size: Mathematics: N = 117 schools, 579 teachers; Reading: N = 117 schools, 629 teachers							

Note. Staffing is not listed as a predictor of Attention to Data because this path does not estimate in either model.

Tables B-2 and B-3 present the covariances among latent variables in the mathematics and reading structural equation models. Covariances measure the level of association between two variables and when standardized range from 0 to 1, similar to correlations where values closer to 1 indicate higher levels of association. In these tables, the associations between latent variables under the same broad category of the theory of action (Supports for Data Use or Working With Data) are presented. The first column describes which covariance is being tested. The next three columns, under the heading "Mathematics," presents the path coefficients (*β*), the standard error, and the associated *p*-value for tests of significant for the specific path/equation tested. Coefficients (*β*) can range in value from 0 to 1, with values closer to 1 indicating a stronger relationship between the variables. Paths are considered statistically significant if the *p*-value is less than 0.05.

Table B-2. Covariances Among Latent Variables in the Mathematics Structural Equation Model

Covariances Among Latent Variables Within Key Dimensions	Coefficient (*β*)	SE	*p*-value
SUPPORTS FOR DATA USE			
Staffing/Human Resources With Infrastructure	–0.08	0.111	0.489
Staffing/Human Resources With Organizational Supports	0.60	0.101	<0.001
Infrastructure With Organizational Supports	0.05	0.089	0.585
WORKING WITH DATA			
Attention to Data With Collaboration	0.68	0.082	<0.001

Table B-3. Covariances Among Latent Variables in the Reading Structural Equation Model

Covariances Among Latent Variables Within Key Dimensions	Coefficient (*β*)	SE	*p*-value
SUPPORTS FOR DATA USE			
Staffing/Human Resources With Infrastructure	–0.09	0.123	0.458
Staffing/Human Resources With Organizational Supports	0.48	0.099	<0.001
Infrastructure With Organizational Supports	0.10	0.082	0.215
WORKING WITH DATA			
Attention to Data With Collaboration	0.72	0.067	<0.001