A Theoretical Framework for Data-Driven Decision Making

Ellen B. Mandinach, Margaret Honey, and Daniel Light
EDC Center for Children and Technology[1]

Paper presented at the annual meeting of AERA, San Francisco
April 9, 2006

[1] The research on which this paper is based has been funded by the National Science Foundation under Grant No. REC 03356653. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. The authors would like to acknowledge the contributions of the staff of the NSF evaluation framework project and other colleagues – Cornelia Brunner, Juliette Heinze, Hannah Nudell, and Luz Ruvas. This paper is based on a presentation made at the Wingspread Conference on Linking Data and Learning, October, 2005.

A Theoretical Framework for Data-Driven Decision Making

Introduction

In the wake of the No Child Left Behind legislation (NCLB, 2001), data-driven decision making has become a central focus of education policy and practice. Schools seek to meet the Adequate Yearly Progress (AYP) requirements of NCLB face tremendous pressure to monitor carefully student performance on the high-stakes assessments that determine their success or failure. The complexity of disaggregating, analyzing, and reporting these testing data has increasingly led administrators to embrace commercial and home-grown data-driven decision making tools and support systems to help track and drive improvement in student performance. Recent work (Stringfield, Wayman, & Yakimowski-Srebnick, 2005; Wayman, Stringfield, & Yakimowski, 2004) notes the rapid proliferation of these tools, and purchasing data further indicates a 17% rate of growth in this market between 2003 and 2004 (Hayes, 2004). One consequence of the increased use of these tools is a growing gap between the use of test data to satisfy administrative demands and the use of test data in concert with other data sources to aid instructional decision making. While these tools have the potential to support the classroom-level instructional decisions of teachers, these tools tend to privilege an approach to data analysis that allows for the examination and reporting of system-wide or school-wide test trends and patterns, but only reveal limited information about individual students and the multiple factors that influence student performance. As a result, they meet the needs of school administrators much more readily than they do those of classroom teachers.

Recent research conducted at the Education Development Center's Center for Children and Technology (EDC/CCT) has found that school administrators use high-stakes test data to understand *general* patterns of performance, identifying class-, grade-, and school-wide strengths and weaknesses so that they can allocate resources and plan professional development and other kinds of targeted intervention activities (e.g., after school remediation, summer school attendance, etc.). Teachers, in contrast, are wary of using any single data source, such as high stakes test data, to make decisions about their students' strengths and weaknesses. Their preference is to engage multiple sources of data – homework assignments, in-class tests, classroom performances, as well as impressionistic, anecdotal, and experiential information – to inform their thinking about student learning (Brunner, Fasca, Heinze, Honey, Light, Mandinach, & Wexler, 2005; Honey, Brunner, Light, Kim, McDermott, Heinze, Breiter, & Mandinach, 2002; Light, Wexler, & Heinze, 2004). While this approach to data yields a richer profile of individual student performance, it also has a downside. Our research and that of others (Confrey & Makar, 2002, 2005; Hammerman, & Rubin, 2002, 2003) suggests that teachers are more inclined to examine factors that contribute to individual patterns of behavior and to think on a case-by-case basis, rather to look for patterns in data at different levels of aggregation, such as classroom-wide patterns. As a result, teachers' decision making strategies often lack systematicity, from student-to-student, class-to-class, and year-to-year, are unintentionally tinged with personal bias, and ignore key statistical concepts like distribution, variation, and reliability.

This paper builds on a project sponsored by the National Science Foundation to explore and create an evaluative framework for data-driven decision making (Mandinach, Honey, Light, Heinze, & Nudell, 2005; Mandinach, Honey, Light, Heinze, & Rivas, 2005). An outgrowth of this work and work of other EDC/CCT projects as well as the that of others at this conference, including the project advisory board, has led us to the development of an emerging conceptual framework for data-driven decision making. One of the goals in this project is to use the data we have collected in six sites across the country to develop, refine, and validate our conceptual model. This work is ongoing and evolving. We will present our current model, couched in the context of data-driven decision making in classrooms, schools, and districts. There is no question that state and federal mandates also influence decision making, as well as many other variables. The larger objective of this project is to capture the dynamic, interactive nature and the complexities of how schools make decisions, across all levels of the districts, across various stakeholders, given the many influences and the contextual surrounds, using systems thinking as an analytical perspective (Mandinach, 2005; Mandinach & Cline, 1994). For the purposes of this paper, the conceptual model, however, focuses solely on the classroom, school, and district levels of decision making, recognizing how the affordances of technology-based tools can facilitate, support, and enable decisions across stakeholders.

The Need for Effective Data-Driven Practices

Research on Systemic Reform and Data Systems

One consequence of the standards and accountability movement is that district and school administrators are being asked to think very differently about educational decision making, and are being asked to use data to inform everything from resource allocation to instructional practice. O'Day (2002) notes the complexity of the mechanisms by which accountability is used in school improvement. Researchers at the UCLA Center for Research on Evaluation, Standards, and Student Testing (CRESST) note that "data-based decision making and use of data for continuous improvement are the operating concepts of the day. School leaders are expected to chart the effectiveness of their strategies and use complex and often conflicting state, district, and local assessments to monitor and assure progress. These new expectations, that schools monitor their efforts to enable all students to achieve, assume that school leaders and teachers are ready and able to use data to understand where students are academically and why, and to establish improvement plans that are targeted, responsive, and flexible" (Mitchell, Lee, & Herman, 2000, p. 22). As CRESST researchers note, "Despite both the mandates and the rhetoric, schools are woefully underprepared to engage in such inquiry. The practice of applying large-scale data to classroom practice is virtually nonexistent" (Herman & Gribbons, 2001, p. 1).

The literature on systemic efforts to improve schools has been focused on the role of data in developing, guiding, and sustaining organizational change that leads to improvements in student learning (Massell, 1998). Initial interest was on data for accountability (Fullan & Stiegelbauer, 1991; Schmoker, 1996), but the debate around measurement driven instruction in the 1980's was an early attempt to use assessment data to improve instructional decision making (Popham, Cruse, Rankin, Sandifer, & Williams, 1985; Shepard, 1991). As is often the case, however, human desires far out paced actual capabilities. For assessment data to be useful for instructional planning they need to be current, accurate, and in the hands of knowledgeable decision makers at the appropriate levels. Yet school systems rarely have had the capacity to process and disseminate data in an efficient and timely manner (Ackley, 2001; Thorn, 2002). Further, advances in school networking infrastructures and online data warehousing have made it feasible to create systems that use assessment data to support decision making, by providing timely information and presentation and analysis tools to educators across multiple levels of the system.

Recently, the education community has again become interested in data-driven instructional decision making, largely because growing numbers of school systems and states have the capacity to process and disseminate data in an efficient and timely manner (Ackley, 2001, Thorn, 2002). This trend has been further accelerated by the requirements of NCLB to use data to improve school performance (Hamilton, Stecher, & Klein, 2002).

The research on data systems and tools to support instructional decisions is a young and emerging field. There is a small and growing body of literature on the use of such data systems, tools, and warehouses to support decision making processes in schools. Stringfield and colleagues (2005) provide one of the first attempts to describe, classify, and evaluate these emerging tools, while Hamilton, and colleagues (2002) offer a brief review of the literature on using test-based accountability data for decision making. As of yet, there is little evaluation across cases of these data tools in application. There are a number of initiatives being implemented across the country for which research is in various stages of development. These projects include the Quality School Portfolio (QSP) developed at CRESST (Mitchell & Lee, 1998), IBM's Reinventing Education initiative in Broward County Florida (Spielvogel, Brunner, Pasnik, Keane, Friedman, Jeffers, John, & Hermos, 2001), the Texas Education Agency and the South Carolina Department of Education (Spielvogel & Pasnik, 1999). There is ongoing work being conducted on data-driven tools in New York, (Brunner, et al., 2005; Honey, 2001; Honey, et al., 2002), Minneapolis (Heistad & Spicuzza, 2003), Boston (Murnane, Sharkey, & Boudett, 2005; Sharkey & Murnane, 2003), Milwaukee (Mason, 2001, 2002; Thorn, 2002; Webb, 2002), and other locations (Chen, Heritage, & Lee, 2005; Lachat & Smith, 2005; Streifer & Schumann, 2005; Wayman, 2005).

Stringfield and colleagues (2005) provide one of the first comprehensive reviews of the tools available, identifying some of the technical and usability issues districts face when selecting a data application to support instructional planning. Technical challenges include data storage, data entry, analysis, and presentation. Other challenges include the quality and interpretation of data, and the relationship between data and instructional practices (Cromey, 2000). Work done on the QSP in Milwaukee indicates that educators are hesitant to base decisions that affect students on data they do not necessarily believe are reliable and accurate (Choppin, 2002). The standardized test data provided in many of these data systems were often not originally intended for diagnostic purposes (Popham, 1999; Schmoker, 2000). Educators' knowledge and training in the use of data is also a confounding factor. While teachers and administrators need not be experts in psychometrics, they must have some level of assessment literacy (Webb, 2002). However, most educators are not trained in testing and measurement and assessment literacy is therefore a major concern (Popham, 1999).

While debate about the merits of using state mandated testing data for diagnostic purposes continues, responding to accountability requirements remains a daily challenge that schools and districts must address now (Pellegrino, Chudowsky, & Glaser, 2001;

Stiggins, 2002). Although high-stakes accountability mandates are not new, the NCLB legislation places public schools under intensified external scrutiny that has real consequences (Fullan, 2000). Not only are failing schools identified, but parents are given the option of removing their children from such schools or using school resources to hire tutors and other forms of educational support. District and school administrators are struggling to respond to these heightened expectations, which by design call for different thinking about the potential of accountability data to inform improvements in teaching and learning. It is clear that NCLB is requiring schools to give new weight to accountability information and to develop intervention strategies that can target the children most in need. The growing interest in data-driven decision making tools is no doubt a direct response to these mounting pressures (Stringfield et al., 2005).

Research in the Service of Practice

The development and implementation of data-driven decision making tools is only one of the necessary steps toward effective use. Several barriers to the effective use of data by educators have been identified (Lim, 2003) including access issues, technical expertise, and training (Choppin, 2002; Cromey, 2000; Mason, 2002; Wayman, 2005). The lack of training for teachers in how to use data to improve student performance has posed a long-term problem (Schafer & Lissitz, 1987; Wise, Lukin, & Roos, 1991). It is rare to find schools in which teachers routinely engage in thinking critically about the relationship between instructional practices and student outcomes (Confrey & Makar, 2005; Hammerman & Rubin, 2002; Kearns & Harvey, 2000). Further, we know very little about the cognitive strategies teachers employ to transform data into useable information and practice (Cizek, 2001; Herman & Gribbons, 2001). And, as noted at the outset of this paper, the kinds of data-driven decision making tools that are proliferating in schools do not provide the kind of detailed data on individual students that could help teachers gather systematic evidence about the effectiveness of particular instructional strategies.

Helping all schools and students achieve, regardless of ethnic and socioeconomic background, requires that we identify and develop processes and practices that support teachers' deep and sustained examination of data in ways that are aligned to local instructional goals. Based on the experience of others (Confrey & Makar, 2005; Hammerman & Rubin, 2002, 2003), teachers need to develop fluency in a number of areas in order to make effective use of data. Teachers do not examine data systematically nor do they make sense of data in terms of long-term trajectories (Confrey & Makar, 2005). They either neglect or fail to understand the concepts of distribution, sampling variation, and statistical difference. Confrey and Makar (2002) find that novice teachers need to know what to look at as they tend to focus on data from individual students, mean scores, and passing rates, while ignoring distributions. The ability to examine a distribution as an aggregate, taking shape of the entire distribution rather than focusing on individual students is a critical skill. But while looking at an aggregate measure such as a mean is more representative of the entire distribution, it runs the risk of ignoring that some students doing poorly are balanced by others doing well, given that the mean is not an indication of variability. Confrey, Makar, and Kazak (2004) found reporting only means and percent passing of high-stakes tests can lead to stereotyping of disaggregated subgroups. Thus, there is a need to understand the concepts of variation and distribution.

When people look at data from a number of different groups all together, they can miss differences among the groups. Understanding what constitutes a significant difference among groups and how to interpret interactions are also critical skills. It is important to define the groups and then examine the distributions of the groups separately to discern potential differences. Understanding that there is normal variability in every process is yet another core skill. Small changes from one time to the next, such as those that occur when retaking the same test, can indicate nothing or something significant. The key is deciding when to pay attention to the differences.

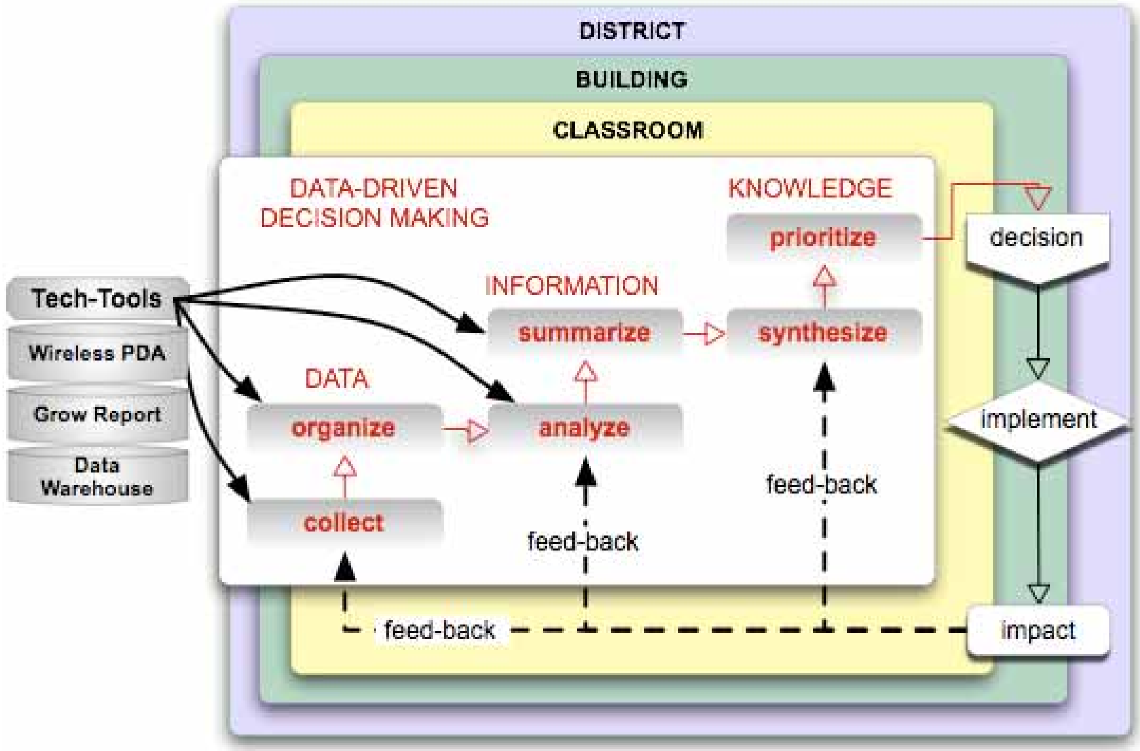
There is no question that data-driven decision making is a complex undertaking, even for the trained educator who understands statistical concepts. As Secada (2001) notes, data should be used to inform educators' decisions, not to replace them, and this process requires time and effort. Educators must have specific uses in mind when examining data, and the decisions they make must be both strategic and timely. Fundamentally, this line of reasoning goes back to the basic definition of validity (Cronbach, 1976; Messick, 1989) which states that validity resides more in test score interpretation than test construction. The hallmark of statistical fluency is understanding how data should be used, the interpretations that can be made from those data, and how such interpretations can be used to guide different types of decisions. For school personnel, a central component in this process is asking good questions about the data, analyzing the data accurately, and then applying the results appropriately (Mason, 2001).

The use of data in meaningful ways assumes at least some level of facility with and knowledge about assessment information. While teachers and administrators need not be experts in psychometrics, they must have some knowledge of general testing concepts. Further, with the proliferation of accountability requirements from the local, state, and federal levels, there is a need for teachers and administrators to be conversant with and make use of the plethora of student assessment data. Fullan (2000) argues that teachers must become assessment literate, and that focusing on student work through assessment is a trait of a good school. According to Fullan and Stiegelbauer (1991), assessment literacy is key step to improve student learning. Further, the appropriate individuals must be armed with the appropriate data. Even if districts have meaningful formative assessments, placing the data in coherent formats is still a major challenge (Ackley, 2001). Thus, the potential power in data-driven tools becomes all the more important for supporting users to collect, analyze, and interrogate data in more effective ways.

Conceptual Framework

The conceptual framework for data-driven decision making is depicted in Figure 1. This framework is founded on the notion of what it means for an educator to be data-driven. We make the assumption here that individuals, regardless of where they are within a school system, have questions, issues, or problems for which data must be collected, analyzed, and examined in order to make informed decisions. This need crosses levels of the organization from the classroom to the school, and to the central administration. As mentioned above, it is important to note that this model presented here depicts decisions made within school districts, focusing on the classroom, building, and district levels. No doubt many variables at the state and local levels can and will impact local decisions, but our intention here is to examine local decisions. The broader range of decisions will be explored through our forthcoming systems-based evaluative framework.

Figure 1. Framework for Data-Driven Decision Making.



The conceptual model we present here has evolved over time and has been informed by the work of colleagues and others. In an initial effort to create a conceptual framework model for data-driven decision making, colleagues at CCT (Light, et al., 2004) have examined organization and management theory in the use of data. In developing a conceptual framework for the use and transformation of data, they drew upon the work of Ackoff (1989), Drucker (1989), and a former CCT visiting scholar (Breiter, 2003). According to Ackoff (1989), data, information, and knowledge form a continuum in which data, are transformed to information, and ultimately to knowledge that can be applied to make decisions. As Light and colleagues (2004) note:

- **"Data** exist in a raw state. They do not have meaning in and of itself, and therefore, can exist in any form, usable or not. Whether or not data become information depends on the understanding of the person looking at the data.
- **Information** is data that is given meaning when connected to a context. It is data used to comprehend and organize our environment, unveiling an understanding of relations between data and context. Alone, however, it does not carry any implications for future action.
- **Knowledge** is the collection of information deemed useful, and eventually used to guide action. Knowledge is created through a sequential process. In relation to test information, the teacher's ability to see connections between students' scores on different item-skills analysis and classroom instruction, and then act on them, represents knowledge." (p. 3)

The continuum provides a logical progression from taking raw data and transforming them into usable knowledge.

Components of the Framework

The data to knowledge continuum provides the foundation for the conceptual framework. It is grounded within the context of the classroom, the school, and the district, all of which will use different data in different ways to make decisions. The role of the technology-based tools is to enable, support, and facilitate decision making by various stakeholders in different parts of the model.

Data-Driven Decision Making Skills

As can be seen in Figure 1, the data to knowledge continuum is defined by the inclusion of six cognitive skills or actions that we have identified as crucial to the decision making process. Two skills are seen to align with each of the points along the continuum. At the data level, the two relevant skills are "collect" and "organize". The skills at the information level are "analyze" and "summarize". At the knowledge level, "synthesize" and "prioritize" are the skills seen as relevant.

A stakeholder, whether a teacher in a classroom, a principal, or a district administrator, is faced with an issue, a question, or a problem for which the *collection* of data should be helpful. The stakeholder must decide what data to collect; that is, decisions must be made about what will inform the issue. The individual then may decide to collect new data or interrogate existing sources of data. For a classroom teacher, this might mean giving students an assignment or activity to highlight a particular learning problem. For a central administrator, it may mean drilling down into the district data warehouse or surveying parents to answer a particular question. Once the data have been collected, it is necessary to *organize* the data in some systematic way so that sense can be made of the data. It is difficult, if not impossible, to extract meaning from raw data that have not been pulled together in some sensible manner. This organizational scheme enables the stakeholder then to convert the raw data into information in which meaning can be imposed.

The stakeholder then takes the organizational scheme created from the raw data and *analyzes* those data for informational purposes. A teacher may analyze results from a classroom exercise. A principal may examine results across classes in a particular grade from a standardized test. A district administrator may analyze trends in performance for various cohorts of students to determine the likelihood of attaining AYP. The scope of the analyses may be broad or constrained, depending on the type of inquiry and the role of the decision maker. Regardless of the depth and breadth, there needs to be some sort of *summarization* of all the accumulated information. Educators are bombarded with information from all directions and from many sources. It is therefore vital to have concise and targeted summaries of information that then can be transformed into usable knowledge, the final stage along the continuum.

To turn information into knowledge, the stakeholder must *synthesize* the available information. The final step is to *prioritize* the knowledge. Setting priorities often requires imparting a value judgment on the accumulated information and knowledge. It necessitates a determination of the relative importance of the information and possible actionable solutions. A teacher may determine that it is more important first to remediate a student's literacy deficit before attempting to address other less pressing learning issues. A principal may determine that it is more important to focus on one curriculum as opposed to another, based on teacher response and student performance. The superintendent may decide that the most potential for solving the minority achievement gap is to allocate resources disproportionately to the most needy schools. Prioritization allows decision makers to determine what is the most important, most pressing, the most prudent, or the most rational solution to a particular educational problem.

The outcome of this six-step process, moving from data to information to knowledge is a *decision*. The decision is then i*mplemented*, or in some instances may fail to be implemented for other external reasons, such as a lack of resources. The implementation results in some sort of outcome or *impact*. Depending upon the impact, the decision maker may decide that he or she needs to return to one of the six cognitive steps, thereby creating a feedback loop. The stakeholder may need to collect more data, may need to reanalyze the information, or resynthesize the knowledge. Because of the feedback loops, data-driven decision making is seen as an iterative process with data leading to a decision, implementation of that decision, determination of the impact, and perhaps the need to work through some or all of the six processes again.

Levels of the School System

The type of decisions and the form of data collected may differ depending on the level at which the decision is being made and by the particular stakeholder. How the data are aligned across the levels of the district will influence the utility. For example, the model of use for accountability will determine how the data factor into the decision making process. That is, the data may be used for facilitation, for progress monitoring, or even for punitive purposes. It is likely that there will be different stakeholders at different levels. It is also likely that there will be different feedback loops (*i.e.*, the iterations within the decision making processes) for different stakeholders and at different levels of the school hierarchy. The feedback loops will depend on the implementation model and context of the data-driven decision making. Data that teachers need will differ from what a building or central administrator may need. The questions asked will differ. Although many questions and the utility of the data may be embedded within a particular level of the school district, there will likely be interactions across the levels. Building level decisions will impact the classroom level just as classroom level decisions will impact the building level. District level decisions will impact the building level and indirectly or indirectly affect what happens in the classroom.

For example, the district administration may decide to introduce a new quarterly assessment for all grades to determine how students are doing in math or in reading, as was instituted on one of our project's districts. This is an example of a top-down decision. Issues that may have impacted the decision might include the resources necessary to carry out such an assessment program, the potential information to be gained, personnel issues around who will develop the assessments and who will score them, and the like. The building personnel then have to carry out the decision. A principal or school curriculum person has to allocate time in the schedule for the administration and scoring of the assessments and determine how the resulting data will be used at the school level to inform instructional and perhaps other decisions. The classroom teachers have to implement the assessments and collect the data once the test results are scored. It is then up to them to use those data to help remediate particular learning deficits for particular students. In this example, the closer to the data a stakeholder is, the more instructional validity the decisions will have due to the proximity to the data and the ability to transform the data into information and then actionable knowledge.

In examining cross-level decision making, it is likely that there will be more top-down decisions than bottom-up decisions. That is, there probably will be fewer decisions made by classroom teachers that will directly influence a decision made at the district level. There will, however, be many within-level decisions. No doubt the cultural surround and context that translate into rationales, needs, and purposes, will determine who uses the data, how they are used, and the sorts of interactions across levels of the stakeholders,

The Role of the Technology-Based Tools

We envision that technology tools can be used to support, enable, and facilitate data-driven decision making. The added value for the use of technology in data-driven decision making is becoming increasingly clear. It is our contention that such tools can be enabling devices for good practice. They have the potential to support data mining that is not possible without technology. But as with many innovations, there is always the "depends" clause when it comes to answering the question about "does technology work" or "what is the impact of the technology".

One issue that staff on the NSF evaluation framework project have debated as we have examined our data over the past two years has been whether data and the tools are separable. Some have argued that the data are embedded within the tools are inextricably linked, and therefore they must be treated together. Others have argued that although the tools influence the data available or that the data influence the tools selected for use, they should be treated as independent entities. We posit that there are complex interactions at play here, recognizing that we function in a multivariate world. Just as there are data by information by needs by value interactions, there also are person by data by tool by context interactions. These interactions relate to the context and values of a school district where decisions are made about the importance of particular kinds of data and the types of technology-based tools that will enable the interrogation of those data.

The affordances of the technology should facilitate the use of the tool and align with the type of data, the contextual surrounds, and the goals for data and tool use. There are characteristics, functions, and capacities of a tool that will facilitate or impede its use. For example, if a tool is too labor intensive, it is likely that individuals will not readily use the application. This project and prior work at EDC/CCT (Light, et al., 2004) have highlighted several functionalities that impact how a tool will be used: (a) accessibility; (b) length of the feedback loop; (c) comprehensibility; (d) flexibility; (e) alignment; and (f) links to instruction.

In terms of these characteristics, accessibility refers to how easy the tools is to access and use. For example, one data warehouse we have examined has a difficult user interface. Regardless of the amount of training, practitioners, particularly at the school level, are impeded by the challenges imposed by the interface. Rather than use the application themselves, individuals have to ask a data person to interrogate the data. Length of the feedback loop refers to the duration between when the data are collected to when the end user receives them in a form that is meaningful. An example here is the difference between two other tools we have studied. Diagnostic handheld devices enable teachers to collect and analyze early childhood literacy or mathematics assessment data and immediately transform those data into actionable knowledge (see Hupert & Heinze, this volume). These handhelds bridge the abyss between assessment and instruction that all too often occurs (Mandinach & Snow, 1999). In contrast, when a standardized test is administered in the spring and the data are delivered in a reporting system to teachers in the autumn, with the intent of using these data to form instructional strategies for particular students, the delay in the feedback loop is substantial and will minimize the utility of the data (Brunner, et al., 2005; Light, et al., 2004; Mandianch, et al., 2005). Thus, the tighter the feedback loop provided by the application and the more recent the data, the more informative the data are likely to be.

Comprehensibility refers to how the tool presents data to the user to facilitate understanding and interpretability. Often a variety of presentation modes improves comprehensibility, such as the inclusion of graphics, tables, and the ability to aggregate the data in different ways. The handhelds again are a good example of providing multiple modes of interpretation. Once data are downloaded to a website, teachers or administrators can view the data in a variety of ways, some graphic, some tabular, and with different levels of aggregation (*i.e.*, the individual student or classroom level results). Similarly, flexibility refers to how a tool allows the user to manipulate the data. In terms of both comprehensibility and flexibility, take for example the two data warehouses the project has examined (Mandinach, et al., 2005). In one instance, the data entered into the warehouse are at an individual student level and there is no provision for the end user to aggregate the data at the classroom, grade, or school levels. If the end user wants such aggregation, a request must be made to the district's research and assessment department who will then make a special data run.[2] In contrast, the other warehouse has the flexibility to aggregate data at multiple levels, enabling the end users to explore the data themselves. The greater the flexibility, the more comprehensible the data become for the even the most novice user.

Alignment refers to how well the tool enables the alignment of the data to the objectives of the stakeholders, making the data useful, meaningful, and of high quality. Scope and accuracy are important factors here. The data must have the depth and breadth to be useful, all the while being vetted as accurate. Both of the data warehouses examined in this project contain the data deemed essential to the districts' goals. Not only do these warehouses contain standardized achievement test scores that determine the districts' AYP status, but also many more data sources that reflect local objectives. The final tool characteristic focuses on links to instruction. The data need to be linked to practice, such as those from the handhelds that can directly inform instruction. Some of the data from the warehouses may be linked to instruction, such as the quarterly writing and mathematics assessments; whereas other data may be much farther removed from instruction, including some standardized achievement tests that may not be fully aligned to standards and instruction.

Given these capacities and functions of tools, a key issue is whether the application enables effective interrogation of the data that match the specified objectives of the end users. This project examined three types of applications that varied markedly across the dimensions of functionality – the handhelds, a paper-and-pencil and online

[2] Based on feedback from our research, the district has begun development to remediate this problem to allow for easier aggregation of data across units of analysis.

reporting system for standardized achievement test results, and the homegrown data warehouses. Other applications are available now or are being developed. What becomes clear is that the selection of a tool is a complex decision based on many factors. The characteristics of the tools will impact how the data are being examined, the types of questions that can be asked, and by whom. The tool also determines the type of data that can be examined and how the data are organized. Conversely, the type of data determines the sort of tool that will support those data. The needs, values, and objectives of the users and the district play a major role in the selection process. Some tools provide access to only student data that are aligned with standards and instruction, while other applications are broader in scope and move well beyond student information management systems (see Wayman, et al., 2004 for a comprehensive review of the possibilities).

Districts need a vision and plan of how data-driven decision making will be implemented and sustained, for what purposes and goals, and by whom. The selection of appropriate tools must align with those objectives. No single tool may be sufficient to meet all the goals. Take for example one of our project sites. We selected this district because they were using the test data reporting system. We soon learned that they also were using the handhelds and developing their own data warehouse. They were using a triangulation strategy, using different tools to meet different information needs. One of our data warehouse sites is considering the adoption of the handhelds to provide immediate diagnostic information. For some districts a single tool will suffice. These decisions must be based on the needs and the resources of each site.

Observations from the Sites: School Issues

We will summarize briefly some of the findings that relate to school issues and impact the framework. These include such factors as accountability and assessment, professional development and training, leadership, and data use.

Accountability pressures by and large are one of the most important factors influencing the use of data and the tools. In the United States, there is increasing pressure at the local, state, and federal levels for schools to achieve performance mandates, as assessed by high-stakes tests. The more tests, the more pressures that are felt by the practitioners, and therefore the need to use data to make informed decisions about instructional practice that my lead to improving achievement, especially given the punitive consequences associated with failure. Because of this increase in testing, schools are faced with an explosion of data. The data need to be mined in different ways, and in particular, must be disaggregated. Simply put, there is so much data that educators are forced to use technological applications to deal with the wealth of data. As many educators say, they are data rich, but information poor. By this they mean that there is far too much information with which they must deal, but those data are not easily translatable into information and actionable knowledge. One goal of using the tools is to facilitate the mining of data from multiple perspectives that ultimately will provide the user with information from which they can make decisions.

A byproduct of the increase in testing is what happens in the classroom in terms of time allocation. As more testing occurs, teachers are forced to devote less time to instruction. Teachers report that they must teach to the tests, and in doing so many important topics get omitted. Teachers feel that they are not teaching as much as they are doing test preparation. Teachers also feel that their typical classroom practices are being changed by these pressures. Teachers know their students and tend to use multiple assessment strategies, quantitative and qualitative, summative and formative, to measure student progress. These strategies translate into a wealth of classroom data that needs to be collected and analyzed. Thus the applications play a critical role in helping educators to manage and examine the plethora of data.

Many teachers feel frustrated by the accountability pressures. Many see the use of data to make informed decisions a necessary survival strategy. Thus the applications by which data can be mined are key tools. Other teachers, however, are taking a more fatalistic approach. They feel that the pressures are just another passing fad and will fade in time, using a strategy to continue practice as usual. While yet another group of teachers are luddites who feel threatened by technology and balk at mining the data, entrusting that task to someone else in their school.

Some teachers' reluctance to use data and the tools is grounded in a lack of training or a mistrust of data. Two kinds of training are salient here. First, there is a need for training on the use and understanding of data. Second, there is the need for appropriate and timely training on the tools. Teachers rarely receive preservice or inservice training. There are relatively few courses offered in teacher training institutions on data, and only recently have such inservice workshops begun to emerge. While teachers need to understand data, they also need to know how to use the technology that makes data mining possible. Again, only recently have professional development opportunities become available.

Leadership is the last major of the major school issues. Leadership makes a difference in terms of the message administrators communicate to their staff. In our experience, building leadership appears to be more important in facilitating or impeding the use of data and the tools. Although superintendents set the tone for a district's philosophy, principles have more direct contact with the faculty and therefore more influence on what they do. A principal who is data-driven or technically savvy can exert substantial influence on the faculty, communicating the importance and thereby stimulating use. In contrast, principals who are luddites communicate that technology and data are not important. They may not be impediments but they certainly do not urge their teachers to make use of the data and technology.

Next Steps

One of the goals in this project is to attempt to validate the conceptual framework from the data we have collected in the six districts. The project was designed so that we have a set of three initial sites where the districts have been implementing one of the three applications, and a secondary set of sites in which we can validate the applications' use. We are analyzing data from these sites to see if the processes conform to the conceptual framework. Already within the course of the past year and a half, the framework has evolved substantially from the initial framework posited by Light and colleagues (2004). Although the data continuum and skills remain somewhat unchanged, the framework presented here acknowledges the impact of the tool, the outcomes of the decision making process, the importance of feedback loops, and the differences among the levels of end users within a district.

Our ultimate goal for this project is to use this conceptual model as the foundation for the developing evaluation framework based on systems thinking promised to NSF

(Mandinach, 2005; Mandinach & Cline, 1994: Senge, Cambron-McCabe, Lucas, Smith, Dutton, & Kleiner, 2000). It will attempt to depict the dynamic nature of data-driven decision making and the complex patterns of interactions among contextual variables, including the impact of the technology-based tools. Fundamental to the systems perspective is that multiple factors influence phenomena through complex interactions and feedback loops. Furthermore, understanding context is essential. Thus, the systems-based evaluative framework recognizes and must capture the dynamic and multivariate nature of phenomena and the hierarchical contextual surround that influences school districts as complex systems

The evaluation framework will extend the conceptual model beyond the level of the school district to include the state and federal levels, acknowledging the need to examine the interconnections within and across levels. More importantly, the evaluation framework will attempt to model the many contextual variables that influence how stakeholders at various levels make decisions. The framework will capture the cultural surround and map the inputs that affect data-driven decision making. The systems approach will enable us to examine how data-driven decision making occurs in educational settings by identifying and understanding the interconnections among components that impact schools as complex and evolving systems.

References

Ackley, D. (2001). Data analysis demystified. *Leadership*, *31*(2), 28-29, 37-38.

Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis, 16*, 3-9.

Breiter, A. (2003). *Information – knowledge - sense-making: A theoretical analysis from management/business literature*. Unpublished manuscript, Bremen, Germany.

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

Chen, E., Hermitage, M., & Lee, J. (2005). Identifying and monitoring students' learning needs with technology. *Journal of Education for Students Placed at Risk*, *10*(3), 309-332.

Choppin, J. (2002, April). *Data use in practice: Examples from the school level.* Paper presented at the American Education Research Association, New Orleans.

Cizek, G. J. (2001). Conjectures on the rise and fall of standards setting: An introduction to context and practice. In G. J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 3-18). Mahwah, NJ: Lawrence Erlbaum Associates.

Confrey, J., & Makar, K. (2002). Developing secondary teachers' statistical inquiry through immersion in high-stakes accountability data. In D. Mewborn, P. Sztajn, & D. White (Eds.), *Proceedings of the twenty-fourth annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education PME-NA24* (pp. 1267-1279), *3*, Athens, GA.

Confrey, J., & Makar, K. (2005). Critiquing and improving data use from high stakes tests: Understanding variation and distribution in relation to equity using dynamic statistics software. In C. Dede, J. P. Honan, & l. C. Peters (Eds.), *Scaling up success: Lessons learned from technology-based educational improvement* (pp. 198-226). San Francisco: Jossey-Bass.

Confrey, J., Makar, K., & Kazak, S. (2004). Undertaking data analysis of student outcomes as professional development for teachers. *ZDM*, *36*(1), 1-9.

Cromey, A. (2000). *Using student assessment data: What can we learn from schools?* Oak Brook: North Central Regional Educational Laboratory.

Cronbach, L. J. (1976). Test validation. In R. L. Thorndike (Ed.), *Educational measurement* (2nd Ed., pp. 443-507). Washington, DC: American Council of Education.

Drucker, P. F. (1989). *The new realities: In government and politics/in economics and business/in society and world view*. New York, NY: Harper & Row.

Fullan, M. (2000). The three stories of education reform. *Phi Delta Kappan 81*(8): 581-584.

Fullan, M., & Stiegelbauer, S. M. (1991). *The new meaning of educational change* (2nd ed.). Toronto/New York, NY: Ontario Institute for Studies in Education: Teachers College Press Teachers College Columbia University.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (2002). *Making sense of test-based accountability in education.* Santa Monica, CA: Rand.

Hammerman, J. K., & Rubin, A. (2002) Visualizing a statistical world, *Hands On!, 25*(2).

Hammerman, J. K., & Rubin, A. (2003). *Reasoning in the presence of variability.* Paper presented at the The Third International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-3), Lincoln, NB.

Hayes, J. (2004, October). *Ed tech trends: A 'first look' at QED's 10th annual technology purchasing forecast 2004-05*. Paper presented at the NSBA's T+L^2 Conference, Denver, CO.

Heistad, D., & Spicuzza, R. (2003, April). *Beyond zip code analyses: What good measurement has to offer and how it can enhance the instructional delivery to all students.* Paper presented at the AERA Conference, Chicago.

Herman, J., & Gribbons, B. (2001). *Lessons learned in using data to support school inquiry and continuous improvement: Final report to the Stuart Foundation* (CSE Technical Report 535). Los Angeles: UCLA Center for the Study of Evaluation.

Honey, M. (2001). *The consortium for technology in the preparation of teachers: Exploring the potential of handheld technology for preservice education*. New York: EDC/Center for Children and Technology.

Honey, M., Brunner, C., Light, D., Kim, C., McDermott, M., Heinze, C., Breiter, A., & Mandinach, E. (2002). *Linking data and learning: The Grow Network study.* New York: EDC/Center for Children and Technology.

Kearns, D. T., & Harvey, J. (2000). *A legacy of learning*. Washington, DC: Brookings Institution Press.

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.

Light, D., Wexler, D., & Henize, J. (2004, April). *How practitioners interpret and link data to instruction: Research findings on New York City Schools' implementation of the Grow Network.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Lim, C. (2003). Data for decisions are just a click away. *ENC Focus*, *10*(1).

Mandinach, E. B. (2005). The development of effective evaluation methods for e-learning: A concept paper and action plan. *Teachers College Record, 107*(8), 1814-1835.

Mandinach, E. B., & Cline, H. F. (1994). *Classroom dynamics: Implementing a technology-based learning environment*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Mandinach, E. B., Honey, M., Light, D., Heinze, C., & Nudell, H. (2005, April). *Data-driven instructional decision making using technology-based tools.* Paper presented at the meeting of the American Educational Research Association, Montreal, Canada.

Mandinach, E. B., Honey, M., Light, D., Heinze, C., & Rivas, R. (2005, June). *Creating an evaluation framework for data-drive instructional decision making*. Paper presented at the meeting of the National Educational Computing Conference, Philadelphia, PA.

Mandinach, E. B., & Snow, R. E. (1999). *Integrating instruction and assessment for classrooms and courses: Programs and prospects for research* (Special Monograph). Princeton, NJ: Educational Testing Service.

Mason, S. A. (2001, spring). Turning data into knowledge: Lessons from six Milwaukee public schools. *Using data for educational decision making. Newsletter of the Comprehensive Center-Region VI*, *6*, 3-6.
Mason, S. (2002). *Turning data into knowledge: Lessons from six Milwaukee public schools* (WP 2002-3). Madison: University of Wisconsin, Wisconsin Center for Education Research.
Massell, D. (1998). *State strategies for building capacity in education: Progress and continuing challenges* (CPRE Research Report RR-41). Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.
Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd Ed, pp. 13-103). New York: Macmillan.
Mitchell, D., & Lee, J., (1998, September). *Quality school portfolio: Reporting on school goals and student achievement*. Paper presented at the CRESST Conference 1998, Los Angeles. Retrieved May 12, 2003, from http://cse.ucla.ed/andmore_set.htm.
Mitchell, D., Lee, J., & Herman, J. (October, 2000). *Computer software systems and using data to support school reform*. Paper prepared for Wingspread Meeting, Technology's Role in Urban School Reform: Achieving Equity and Quality. Sponsored by the Joyce and Johnston Foundations. New York: EDC Center for Children and Technology.
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.

No Child Left Behind Act of 2001. (2001). Retrieved January 30, 2002, from http://www.nclb.gov.
O'Day, J. A. (2002). Complexity, accountability, and school improvement. *Harvard Education Review*, *72*, 1-31.
Pellegrino, J.W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press
Popham, W. J. (1999). "Why standardized tests don't measure educational quality." *Educational Leadership 56*(6), 8-15.
Popham, W. J., Cruse, K. L., Rankin, S. C., Sandifer, P. D., & Williams, R. L. (1985). Measurement-driven instruction. *Phi Delta Kappan, 66*, 628-634.
Schafer, W. D., & Lissitz, R. W. (1987). Measurement training for school personnel: Recommendations and reality. *Journal of Teacher Education*, *38*(3), 57-63.
Schmoker, M. J. (1996). *Results: The key to continuous school improvement*. Alexandria, VA: Association for Supervision and Curriculum Development.
Schmoker, M. (2000). The results we want. *Educational Leadership*, *57*(5), 62-65.
Secada, W. (2001, Spring). From the director. *Using data for educational decision making. Newsletter of the Comprehensive Center-Region VI, 6,* 1-2.
Senge, P., Cambron-McCabe, N., Lucas, T., Smith, B., Dutton, J., & Kleiner, A. (2000). *Schools that learn*. New York: Doubleday.
Sharkey, N., & Murnane, R. (2003, April). *Helping K-12 educations learn from assessment data*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Shepard, L. (1991). *Will national tests improve student learning?* (CSE Technical Report 342). Los Angeles: Center for the Study of Evaluation, UCLA.

Spielvogel, R., Brunner, C., Pasnik, S., Keane, J. T., Friedman, W., Jeffers, L., John, K., & Hermos, J. (2001). *IBM's reinventing education grant partnership initiative: Individual site reports.* New York: EDC/Center for Children and Technology.

Spielvogel, B., & Pasnik, S. (1999). *From the school room to the state house: Data warehouse solutions for informed decision making in education*. New York: EDC/Center for Children and Technology.

Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*. *83* (10), 758-765.

Streifer, P. A., & Schumann, J. A. (2005). Using data mining to identify actionable information: Breaking new ground in data-driven decision making. *Journal of Education for Students Placed at Risk*, *10*(3), 281-293.

Stringfield, S., Wayman, J. C., & Yakimowski-Srebnick, M. E. (2005). Scaling up data use in classrooms, schools, and districts. In C. Dede, J. P. Honan, & L. C. Peters (Eds.), *Scaling up success: Lessons learned from technology-based educational improvement* (pp. 133-152). San Francisco: Jossey-Bass.

Thorn, C. (2002). *Data use in the classroom: The Challenges of implementing data-based decision making at the school level*. Madison, WI: University of Wisconsin, Wisconsin Center for Education Research.

Wayman, J. C., Stringfield, S., & Yakimowski, M. (2004). *Software enabling school improvement through analysis of student data* (Report No. 67). Baltimore: Johns Hopkins University, Center for Research on the Education of Students Placed at Risk. Retrieved January 30, 2004, from www.csos.jhu.edu/crespar/techReports/Report67.pdf.

Wayman, J. C. (2005). Involving teachers in data-driven decision making: Using computer data systems to support teacher inquiry and reflection. *Journal of Education for Students Placed at Risk*, *10*(3), 295-308.

Webb, N. (2002, April). Assessment literacy in a standards-based urban education setting. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Wise, S. L., Lukin, L. E., & Roos, L. L. (1991). Teacher beliefs about training in testing and measurement. *Journal of Teacher Education*, *42*(1), 37-42.