



Center for Education Policy Research
HARVARD UNIVERSITY

Balanced Scorecards and Management Data

Working Paper Series: February 2009

FREDERICK M. HESS and JON FULLERTON

The authors would like to thank Thomas Gift for his invaluable research and editorial assistance and Michael Eugene, Erin McGoldrick, and Bryan Richardson for their feedback and suggestions.

Balanced Scorecards and Management Data

FREDERICK M. HESS is director of education policy studies at the American Enterprise Institute. JON FULLERTON is executive director of the Center for Education Policy Research at Harvard University.

Successful organizations, public and private, monitor their operations—extensively and intensively. UPS and FedEx know where every package is in transit. Dell is famous for running an extremely tight supply chain, pushing the cost of holding inventory onto its suppliers by having a crystal clear understanding of its immediate requirements and only ordering what it needs when it needs it. Baseball teams employ sophisticated statistical analyses in making personnel decisions.

Compare such approaches to what has long prevailed in public education. In 2007, Michelle Rhee, then the new chancellor of the Washington, D.C. Public Schools, reported that millions of dollars worth of textbooks and supplies had been moldering, unnoticed, in a warehouse for months and years. Few districts understand their true costs of recruiting a new teacher and principals have little idea what their schools' actual budgets are.

One consequence of this data drought is that school systems focus single-mindedly, even obsessively, on the few metrics they do have—such as test scores and expenditures. Even districts that tout themselves as “data-driven” often mean only that they can break down test scores by teacher, subject, or student population; few, in our experience, have reliable information on how satisfied principals are with the support provided by human resources or how rapidly the information technology team addresses instructional requests. Generally speaking, this is as true when it comes to systems of charter schools as it is for traditional districts. This dearth of data makes it difficult to manage and improve the critical functions that support teaching and learning. For instance, many urban systems, though desperate for talent, are unresponsive to inquiries from promising candidates. However, absent good data on this count, senior officials are rarely even attuned to the problem.

The New Teacher Project's (TNTP) 2005 study *Unintended Consequences* provides a compelling illustration of how good management data can change this equation. The study reported that transfer and excess rules forced many urban schools to hire teachers they did not want while preventing them from removing teachers deemed unsuitable. On average, 40 percent of school-level

vacancies were filled by voluntary transfers or “excessed” teachers in which schools had either no choice at all or limited choice in hiring. TNTP collected data from labor relations staff and reported that districts typically terminated only one or two tenured teachers a year for inadequate performance.¹ In 2005, prodded by the furor surrounding *Unintended Consequences*, the New York City Department of Education and the United Federation of Teachers signed a landmark contract that reformed the staffing process for teachers and schools by enabling schools to choose which teachers they hired, regardless of seniority; ending the “bumping” of novice teachers by senior teachers; and increasing transparency in hiring. In 2008, TNTP reported that, in the first two hiring seasons, the new system allowed over 7,500 transfer and excessed teachers to obtain jobs at new schools, with 90 percent of transfer teachers and 80 percent of excessed teachers describing their new placements as satisfactory.²

Put plainly, it is difficult to manage modern organizations for breakthrough improvement without accurate, timely data and the knowledge and willingness to use them. Yet we see a vacuum in schooling when it comes to collecting crucial data that stretch beyond reading and math scores and auditable enrollment and financial information. Test results are an important measure of student learning. Attendance is an important element of budgeting. But ensuring the high-quality provision of services requires operational measures and data well beyond those of student achievement and body counts.

Districts need to complement such basic data with reliable measures that illuminate the performance of complex operations like human resources, procurement, and data management, at both the district and school levels. Developing and tracking appropriate metrics is the starting point in enabling effective management. Data should primarily measure not compliance (e.g., was a regulation followed?) or inputs (e.g., how much was spent?) but the efficiency, effectiveness, and quality of district services (e.g., the cost of recruiting a new math teacher, the percentage of textbooks distributed on time to the proper schools and classrooms, or how rapidly teachers can access assessment data).

While discussion of “data-driven” schooling revolves today around the narrow tasks of identifying effective teachers and students who need added assistance, managing with data is more broadly concerned with making schools and the school system more supportive of teaching and learning. Doing so requires tracking an array of indicators, including the shipment and distribution of books and materials and the satisfaction of teachers with the results; the speed at which maintenance workers address school-level concerns; the percentage of teachers who rate the professional development they receive as helpful; and turnaround time on assessment data and the frequency with which those data are employed by teachers. A school system which has these kinds of data is one where management is equipped to revolutionize how schools work, how teachers are supported, and how dollars are spent.

¹ Levin, Jessica, Jennifer Mulhern, and Joan Schunck. *Unintended Consequences: The Case for Reforming the Staffing Rules in Urban Teachers Unions Contracts*. The New Teacher Project, 2005.

<http://www.tntp.org/files/UnintendedConsequences.pdf>

² Daly, Timothy, David Keeling, Rachel Grainger and Adele Grundies. *Mutual Benefits: New York City's Shift to Mutual Consent in Teaching Hiring*. The New Teacher Project 2008.

http://www.tntp.org/files/TNTP_Mutual_Benefits_042708F.pdf

WHY ACHIEVEMENT DATA AREN'T ENOUGH

Over the past ten years, there has been for the first time a concerted push to hold schools accountable for their results by looking principally at student achievement data. Accountability efforts—and particularly the 800-pound-gorilla of No Child Left Behind-style testing—have created an appetite for data. Districts are collecting more achievement data than ever before, and states and districts are becoming less and less diffident about holding schools accountable for results. Many think we are on the verge of a management revolution in using data to drive achievement.

In practice, however, there is a rarely acknowledged tension between collecting data with an eye to public policy and external accountability (measurement *of* performance) and doing so for purposes of internal management (measurement *for* performance). The data most useful to parents and policymakers are often straightforward data on how well students and schools are faring on content assessments; whereas the key data for district officials seeking to facilitate improvement are data that shed light inside the “black box” of the school and district—illuminating *why* those results look like they do. This is why the public financial reports by corporations like General Electric or Google are quite different from the measures that managers there use when seeking to improve operations or practices.

Most available achievement data are of limited use for management purposes. First, state testing regimes tend to provide measures of achievement too coarse to be of much use to teachers and principals seeking to change practice. Those districts serious about using data have adopted benchmark and/or formative assessment processes to supplement state tests. Second, a surprising number of districts are unable to easily link teachers and students in their student information systems. There can be little incentive to address this challenge, given substantial political resistance to such linkages from some teachers, unions, and others who are concerned that such data systems will be used to evaluate teachers on the basis of student achievement.³ Finally, while achievement tests are a useful measure of educational outcomes, they do not capture all that we expect from schools. We expect schools to teach subjects (art, music) and skills (cooperation, self-discipline) outside the reach of appropriate and cost-effective testing regimes.

Even if these issues with achievement data were resolved, there are three problems with focusing solely on such outcomes. First, student achievement measures are largely irrelevant to motivating and managing a large number of important district employees. Does it really make sense to hold a payroll processor responsible for student achievement rather than the speed and accuracy of his/her work? Or for the percentage of principals and teachers who rate the payroll office’s service as courteous and attentive? In fact, it is not clear that it makes sense to encourage districts to evaluate trainers, recruiters, or data analysts on student test scores rather than on indicators which more precisely measure the quality of their work. By focusing so relentlessly on achievement, especially in just a few skill domains, many employees are either excused from results-driven accountability or held accountable for things over which they have little control. The result of this is to undermine the development of a performance ethic and foster cynicism.

Second, it is easy for even talented educators to give short shrift to the operations, hiring, and financial practices that can support educators in schools and classrooms. Operations are like the

³ In fact, the New York State legislature recently included language in its budget bill putting a moratorium on using student performance data for granting tenure to teachers. See Hakim, Danny and Jeremy W. Peters. “Legislators Balk at Tying Teacher Tenure to Student Tests.” *New York Times*, April 9, 2008.

air we breathe in that we scarcely notice the air around us until something goes awry, at which point there can be devastating results. Focusing on “instructional leadership” is difficult when the hiring process does not value teacher quality and assigns instructors to schools with little time to prepare for the new academic year, or when principals and teachers must wait weeks or months for assessment results. Management must monitor the overall effectiveness of key operations, as well as how those operations translate to the school level.

Finally, student achievement data alone can only yield a “black box.” They will not allow organizations to *diagnose* problems and *manage* improvement. If math scores are disappointing, why is that? Is professional development the problem? Is hiring? It is as if a CEO’s management dashboard consisted of one item—the stock price. In fact, given the state of most student achievement data systems, the better analogy is to *last year’s* stock price.

District management needs to create the preconditions and processes that foster high achievement; doing so, however, requires metrics and data that stretch well beyond student achievement. Ultimately, education leaders need to take a page from the “balanced scorecard” approach that has reshaped how private and public sector firms have approached data and management.⁴ Developed in the early 1990s by Robert Kaplan and David Norton, the balanced scorecard seeks to provide a quick but comprehensive view of firm performance. It includes standard financial metrics that reflect past performance but, crucially, complement these with operational metrics on customer satisfaction, internal processes, and the organization’s learning and innovation capabilities—the key predictors of *future* success.

In 1992, Kaplan and Norton explained in the *Harvard Business Review*, “Managers should not have to choose between financial and operational metrics. In observing and working with many companies, we have found that senior executives do not rely on one set of measures to the exclusion of the other. They realize that no single measure can provide a clear performance target or focus attention on the critical areas of the business.”⁵

The balanced scorecard, which by 1998 had already been adopted by an estimated 60 percent of U.S. Fortune 500 companies, was a response to the recognition that relying solely on financial metrics could create distortions.⁶ An emphasis on short-term financial numbers can readily lead firms to sacrifice long-term viability. The proponents of the balanced scorecard approach recognized that enormous value resided in hard-to-measure areas like customer relations, information technology, and employee skills. They realized that effective management required collecting and monitoring performance and cost data on a range of activities that go beyond the “bottom line.”

Well-designed balanced scorecards develop a clear link between operational metrics and financial performance. They install long-term financial performance as the primary objective and then identify how various operational metrics impact that outcome. Ideally, the balanced scorecard brings together, in a single management tool, many ostensibly disparate corporate concerns, such as improving customer relations, boosting product quality, investing in research and development, and developing employees.⁷

⁴ Kaplan, Robert S. and David P. Norton. *The Balanced Scorecard: Translating Strategy into Action*, Boston: Harvard Business School Press, 1996.

⁵ Kaplan, Robert S. and David P. Norton. “The Balanced Scorecard- Measures that Drive Performance.” *Harvard Business Review*, Jan.-Feb. 1992.

⁶ Silk, Scott. “Automating the Balanced Scorecard.” *Management Accounting*, Vol. 11, No. 17, 1998.

⁷ Kaplan, Robert S. and David P. Norton. “The Balanced Scorecard- Measures that Drive Performance.” *Harvard Business Review*, Jan.-Feb. 1992.

In education, employing the balanced scorecard entails articulating goals for student achievement and other key student outcomes (such as completion rates) and then translating them into measures for improving operational efficiency inside and outside the classroom.

LEVELS OF SOPHISTICATION IN DATA COLLECTION

While most districts do not yet assemble the kind of data managers need, districts already collect much more than student achievement data. The amount of financial reporting alone that state and the federal governments require for compliance purposes is absurdly extensive. Indeed, these state and federal demands have historically resulted in data collection that monitors enrollment and minutely tracks broad program and personnel costs. Given limited manpower and expertise, and dated computer systems, district officials will privately concede that they have emphasized these auditing exercises rather than potentially more useful management metrics.

The kinds of changes necessary to turn school systems into high-performing organizations will be dramatic. Even districts routinely heralded as data-driven and high-performing have often not invested in the technology, hired the personnel, or developed the requisite expectations, feedback loops, processes, and analytic competencies. Consequently, many schools and systems are today at the very edge of their capacities when they seek to produce student-level achievement data in a timely fashion in order to ensure that teachers can put that data to work.

We do not term a hospital “well-run” because its doctors make proper use of diagnostic tools; instead, we would reserve that label for hospitals where staff are competent and efficient, supplies carefully tracked and promptly refilled, data files up-to-date, personnel needs quickly handled, and the facility well-maintained. Yet, in schooling, systems that have embraced only the most fundamental elements of professional practice are heralded (and regard themselves) as paragons of modern management.

What would it take for school systems to start collecting the data that would make possible breakthrough management? There are six key needs, forming a rough hierarchy.

1. Accurate collection of basic student, financial, and human resource data: The first step is for any organization to collect the most fundamental data on what it does and how it spends its money. School systems are generally pretty good at this. Federal law now requires school systems to test students and collect basic achievement and attainment data. Basic financial management requires districts to ensure that accounts are not overspent, that school enrollment and attendance figures are accurate, and that only authorized positions are on the payroll. Intergovernmental grants (such as Title I) require that districts account accurately for how they spent the money received and show that it was spent in accordance with regulations. Most districts are already well along on this count, as any district not doing this effectively will run into legal and financial trouble.

2. Data linked over time: Once districts have the initial building blocks, the key is to link them across time. This is essential if leaders are to determine how to improve performance. In general, a district that can collect its basic data accurately can also link them longitudinally. However, there are significant exceptions. Some systems do not maintain consistent identifiers across years for students or employees. One common problem is that organizational change is often not accounted for in financial coding systems. Districts may assign costs only to offices (such as the office of instruction) and not to functions (such as math professional development). The result is

that when a district reshuffles its organizational chart (not an uncommon occurrence!) and math professional development is reassigned to human resources or a new office, it becomes impossible to make comparisons over time.

3. Customer service and satisfaction data: Every company knows that its existence depends upon the satisfaction of its customers. Great companies measure customer service from several dimensions (internal and external) to quickly diagnose operational or professional issues that will hurt their ability to maintain the confidence of the people they serve. While many district and school officials may seek piecemeal information on the satisfaction of employees or parents, these efforts tend to be haphazard. Real progress on this count requires that customer service and satisfaction data be routinely and systematically collected and analyzed.

4. Data with sufficient granularity to illuminate units and activities within departments: Measuring efficiency requires capturing outputs as well as practices and processes that otherwise remain in the vague cloud called “overhead.” For example, when considering the role of human resources, there are various metrics that might help illuminate how resources are being used and opportunities for improved productivity. One set would assess how long it takes a human resources department to vet, interview, consider, and hire or reject an applicant. Others would reflect how human resources managers apportion their time, such as how much time is devoted to engaging in various kinds of recruitment efforts, to addressing the concerns of existing employees, or to handling workers’ compensation. It is the exceptional district that collects that sort of data or monitors them in a fashion that permits useful analysis. Typically, systems will know how much is spent on human resources and the size of the staff, but not how much time the human resources staff spends on recruitment or responding to the needs of teachers and principals. This is a key step in the journey from basic data to useful management data.

5. Data connected across content areas (and to outcomes): Even if the efficiency of human resources processes has improved and vacancies are filled more rapidly, more is needed to judge human resources’ effectiveness. Do the new teachers achieve better or worse student outcomes than the teachers that came before them? Do they stay longer? Are they more or less satisfied with the district’s support services? What about the new principals? Do they “lead” better? Do students in their schools learn more than students in other schools? What would be the financial impact of adding new human resources staff? What would be the expected improvement in processing time or yield? Answering these questions requires connecting human resources system data to student-level longitudinal test data to retention data to survey data. Similar connections are necessary to examine the efficacy of professional development (e.g., which teachers get what services—and do they matter?) and student interventions (e.g., does a pullout program work to improve student achievement?). With this level of data sophistication, it becomes feasible to start conducting cost-benefit analyses of programs, services and organizational units.

6. Doing the above in real time: Ideally, district management should be able to find out instantly which schools are waiting for textbooks or which teachers have received what professional development. While FedEx can tell millions of customers exactly where their packages are around the world, large school systems routinely lose track of thousands of dollars worth of textbooks and supplies.

When districts can marry information on operations and activities to particular educational or intermediate outcomes, they enable managers to gauge relative program effectiveness. When all the pieces are in place, it becomes possible to engage in meaningful cost-benefit analysis. This

would permit a district to know not only the relative costs for each teacher recruited by The New Teacher Project rather than its own human resources operation, but also the relative effectiveness of teachers coming from each route—allowing an evidence-based judgment about the value of alternatives.

Few or no school systems have all of these elements in place today. Most are currently at step two. Consultants or internal district analysts can, with enough time, manpower, and supplemental data collection, provide school systems with analyses that may push to steps four and five. The challenge is for districts to consistently reach step six.

THE NUMBERS WE NEED

So what kinds of data should school systems be collecting and connecting? There are six major strands deserving attention. Unfortunately, even those that have been an ostensible priority have been shortchanged by a tendency to focus less on what will help managers improve schools and systems than on what elected officials need to police fiscal improprieties or measure school performance.

The first and most important, type of data to collect is student outcomes. Just a decade ago, most districts had abysmal systems for tracking pupil achievement and school completion. Today, too many problems still exist, but most school systems can provide coherent data on how well students are doing on state assessments. However, outcome metrics beyond state assessments are often difficult for management to come by. Key data in this field include:

- Performance of students on various sub-strands (e.g., number sense, spatial relations on the math test) of state tests with results taken down to (and accessible to) the classroom teacher.
- Item-level analysis at the individual student and classroom level. This allows teachers to analyze whether all or most of their students miss the same test items—and then to adjust their teaching strategies.
- Results of benchmark tests provided back in a timely manner (e.g., no more than one or two days after the test is completed).
- Employment or enrollment status of students after high school.

The second domain is that of counting and tracking people and things. Monitoring the number of students and teachers, the state of facilities, and relevant district assets are all necessary to provide operational baselines. School systems have historically been good at tracking these kinds of data, largely because state and federal requirements led districts to configure their data systems accordingly. Unfortunately, there has been much less effort at ensuring that these descriptive data are captured with sufficient granularity (as individuals rather than as broad categories) or that they can be matched with expenditures, programs, and outcomes. Key elements would include:

- Authorized staff positions, the location of the positions, the purpose and reporting relationships of the positions, whether they are filled and by whom, and whether they are full or part time.
- District assets and materials, where they are located, and the transfer of assets between locations (e.g., the delivery of textbooks).
- Students, which schools and classrooms they attend, and the teachers and staff in those schools and classrooms. This should include not just the “teacher of record” for the students, but also

aides, tutors and other staff working with the student.

- Teacher and student attendance—and the reasons for absences.

When it comes to finance, systems have invested great effort in developing a capacity to keep track of transactions but little in tracking expenditures in ways that facilitate useful analysis. Developing a management-friendly system for tracking expenditures would require ensuring that managers can link dollars and time spent by employees to locations, activities, reporting structures, and, if appropriate, students. If a professional development coach or a gifted-and-talented teacher works at multiple locations, this should be reflected in financial and payroll data and linked to the teachers and students in question so that the cost-effectiveness of the activity can be monitored and assessed. Some key elements that are often not tracked well include:

- Are dollars actually being spent in specific schools and classrooms or are they being spent by central administration and then “allocated” to school sites based on calculations and projections (e.g., total heating costs for the district distributed proportionally to all schools by number of students)? For instance, schools could be charged per teacher for the average teacher salary cost for the whole district, or schools could be charged the actual salaries of the teachers working at the site.
- Who controls the decision to make the expenditure and for whom does the expenditure take place? For instance, is a school-based professional development program purchased by the office of instruction at the central office or by an individual principal or by an individual teacher? Each of these expenditures are for teachers at the school; however, those held accountable for these expenditures should be quite different.
- What program, activity, and function do the expenditure support?

Fourth, while attention to “instructional leadership” and “capacity building” has led the current generation of district leaders to devote increased time to providing professional development and related resources, few districts track instructional and curricular services in a manner that makes it possible to determine who got what services when. As a result, district leaders are unable to identify particularly effective tactics or programs, effective or ineffective personnel, points of concern, or opportunities for cost savings. Key data on instructional and curricular services include:

- What professional development is delivered to which personnel, when, for what length of time, and by whom?
- What tutoring or afterschool programs are delivered to which students, when, for what length of time, and by whom?
- Which reading programs and which math programs are being used by which schools? How well are they implemented, at what cost, and with what results?
- What texts and ancillary materials are utilized by which schools, classrooms and students?

Fifth, more crucial than any other element of school system management may be human capital operations. Dramatically improving the quality of teaching and learning requires that a school system be able to monitor personnel; to gauge performance; to compensate or remediate in response to performance; and to manage competently hiring, transfer, benefits, employee concerns, and termination. The key is to measure human capital operations not in terms of inputs (number of hires or percentage of educators with advanced degrees) but with metrics that reflect meaningful performance. Key data on human capital include:

- The quality of new hires, in terms of skills, experience, past performance, qualifications, or interview grades.
- The quantity of applicants for positions, how rapidly they are screened and offers made, and the rapidity with which successful applicants are placed and prepared.
- The satisfaction of employees with the support and responsiveness of human resources to various concerns.
- The performance of instructional personnel, support staff, and school leaders as measured by student progress (potentially including standardized assessments, promotion, graduation, course selections, and attendance).
- The performance of personnel on relevant metrics beyond student achievement (e.g., soliciting “forced rankings” of teachers by their principals or supervisors, while systematically collecting evaluations of supervisors by their staff).

Finally, it is essential to monitor business practices like procurement, information technology, data management, and maintenance which facilitate system operation. The functioning of these elements is crucial to effectively support school leaders, classroom educators, and school communities. The key, again, is to measure these services not in terms of inputs but in terms of core metrics that accurately reflect performance:

- How long does it take the district to process a supply request, how rapidly are supplies delivered to the classroom, and how does the system’s cost per order compare to benchmarks?
- How rapidly are school personnel able to access the results of formative assessments, how satisfied are they with the user-friendliness of the data interface, and how intensively/extensively do faculty make use of formative assessments and student data?
- How rapidly does the facilities team respond to complaints and what percentage of complaints is resolved on the first visit? How many work orders do maintenance teams perform in a week?
- What is the cost per square foot of maintenance and what is the staff satisfaction rate with the physical condition of the school?

THE POWER OF DATA

Collecting, maintaining, and employing these kinds of information will permit school and district leaders to manage in profoundly different ways. They will make it possible for them to help professionals fully utilize their skills; eliminate unnecessary or redundant tasks, programs, and personnel; and target resources and effort more effectively.

How might this work in practice? One illustration is provided by the remarkable success that New York City and other cities enjoyed using new data tools to combat crime in the 1990s. The New York Police Department’s system, Compstat, short for “computer statistics,” compiled data from street cop reports, crime complaints, arrest and summons activities, crime patterns, and police activities and used this information to help target police patrols. Over time, the system was broadened to include 734 categories of concern, including the incidence of loud parties.⁸ Compstat made it easier to hold officers accountable, to pinpoint areas of concern, and to

⁸ Anderson, David C. “Crime Control by the Numbers.” *Ford Foundation Report*, Winter 2001. See also Rashbaum, William K. “Crime-Fighting by Computer: Scope Widens.” *New York Times*, March 24, 2002.

provide real-time data to assist both officers and street cops in making decisions. Precincts were required to update crime statistics on a daily or weekly basis, rather than on the monthly or quarterly basis that had been the norm. New mapping software allowed department officials to identify crime clusters by neighborhood and then correlate them with drug sale sites, addresses of known felons, areas of gang activity, and public housing—and to communicate all this information department-wide within seconds.

In the first five years after the 1993 introduction of Compstat, the number of homicides in New York City fell from 1,946 to 629—a rate of decrease three times that of the nation as a whole. In Philadelphia, Compstat was implemented in 1998. In the first year, the murder rate and auto theft rate both fell by more than 15 percent. Similar results were experienced in other cities, including Los Angeles, New Orleans, Albuquerque, Sacramento, and Omaha.⁹

The system worked equally well in other domains. When the New York City police extended the system to traffic control in 1998, vehicle accidents fell 38 percent and pedestrian fatalities declined 36 percent in the first six months. These improvements were credited to the system's ability to highlight the need for small changes like fixing a stop sign, changing light timing, and placing orange nylon mesh at intersections to prevent pedestrians from stepping too far into the street.¹⁰

The Council of Great City Schools has recently begun a comprehensive benchmarking process across a whole series of “meat and potatoes” metrics for business operations such as transportation costs per student, food services participation rates, and lead time required for procurement.¹¹ This is the first time these types of data have been collected for school systems, and their power is evident. Michael Eugene, business manager for the Los Angeles Unified School District and a driving force behind the benchmarking project, has explained the importance of comparative statistics on outcomes:

“I didn't know we had one of the lowest meal participation rates among secondary students until I saw the benchmark data. Between 2002 and 2006 we improved from 17 percent of secondary ADA [Average Daily Attendance] to 37 percent participating in the lunch program, so I thought we'd improved significantly based on trending ourselves over time. But when I saw where we were in the benchmark data my heart sank. Still being among the lowest in the nation blew me away. It has created the utmost urgency to break down barriers of access to nutrition. Until 2002, food services was measured by its profitability rather than its participation rate. While fund balance is important, clearly the District was focusing on the wrong KPI [Key Performance Indicator].”¹²

One little noted, but very important, benefit of Compstat, benchmarking and other such processes is that they force managers to make sure the data are accurate. A lot of bad data are stored and analyzed because middle managers in the organization neither use nor are held accountable for data accuracy. Once attendance rates or dropout rates of individual schools are benchmarked, officials have much more incentive to ensure that the numbers are correct.¹³

⁹ Dussault, Raymond. “Maps and Management: Compstat Evolves.” *Government Technology*, April 2000.

¹⁰ Marzulli, John. “Cops Tackle Road Perils: Computer Helping to Identify Hotspots.” *New York Times*, July 19, 1998.

¹¹ *Managing for Results in America's Great City Schools: A Report of the Performance Measurement and Benchmarking Project*. Council of the Great City Schools, April, 2007.

¹² Eugene, Michael. Personal communication, May 2008.

¹³ Of course, managers also have to be aware of the possibility of audit in order to prevent “over-counting.”

WHAT'S THE PROBLEM?

Everything we have said so far seems pretty obvious and is the way that almost any large, well-functioning organization operates in the 21st century. Why, then, is the collection and analysis of basic student achievement data and so little else regarded as the cutting edge when it comes to managing with data in schooling? Political, cultural, and organizational tensions explain the current paucity of important management data in K-12 education. Five deserve particular mention.

First, and most significantly, our school systems do not reward educational leaders for pursuing new efficiencies, redeploying resources, or coming up with innovative delivery mechanisms for school services. Indeed, superintendents or principals who use management data to propose the elimination of redundant personnel or to zero out ineffective programs are likely to ignite firestorms and invite political conflict. Even if successful, leaders are typically not rewarded (professionally, monetarily, or otherwise) for such decisions. School leadership as a whole is a highly political field, one where a reputation for consensus-building and peacemaking is a treasured asset. So long as the aggressive use of management data is not rewarded, there is little mystery as to why it is rarely collected or employed.

Similarly, because state and federal statutes, salary structures, and existing commitments mean district and school officials have a limited ability to redeploy resources, there is not a lot of incentive to collect data whose value is their ability to steer such determinations. District and school leaders often feel more like overseers of a formula-driven budget than like active participants in shaping those budgets.¹⁴ The result is a chicken-and-egg situation, in which districts have limited incentive to assemble these data, because they have only limited ability to use it, yet the data vacuum that results makes it more difficult to argue to policymakers that new flexibility will be utilized in informed and appropriate ways. This dilemma makes clear that discussions about data and statistics must proceed in tandem with broader policy proposals.

Second, public education has underinvested in its information technology infrastructure for years. The problem is that updating such infrastructures is expensive in the short run—both in terms of dollars and political capital. When a superintendent is faced with the choice between spending millions on information technology or “putting that money into the classroom,” few will opt to explain to parents, teachers, or school board members why they are putting money into data systems rather than class-size reduction, pay raises, or art programs. In the private sector, management can justify such investments by pointing to the bottom line—such an approach, even when compelling, is a more difficult pitch for educational leaders.

Moreover, as recent implementations of new payroll and planning systems in Chicago and Los Angeles show, there are undeniable risks to major upgrades in such systems. Installing a new financial or human resources system is a complex undertaking that often requires employees to change routines and that is challenging even for high-functioning organizations. Even when such installations are successful, design, procurement, implementation, and training mean that the results will not be manifest for several years, while the headaches and costs emerge in the short term. Moreover, if not managed carefully, the result of these efforts can be disastrous—especially in a sector where these efforts are so rare that there's limited expertise on the part of either vendors or peer districts. Los Angeles Unified School District (LAUSD) has spent over

¹⁴ Indeed, even if cost efficiencies are discovered and executed, given the politics of public budgeting processes, any savings are likely to be driven into covering existing deficits or routine increases in employee salary and benefits, not strategic change.

a year sorting out problems from the introduction of its new integrated financial and business operations system. The local union established a permanent RV camp outside district headquarters to highlight mistakes in the payment of teachers, and has used this to indict LAUSD management more broadly. Given that plenty of private sector installations of such systems also have significant difficulties (in fact, one survey suggested that almost half of such Enterprise Resource Planning installations are deemed unsuccessful¹⁵), who can blame superintendents for not wanting to take on such projects. If they succeed, few will notice; if they fail, the costs are high.

Third, while “data-driven” instruction has become a popular buzzword, the cultures of school districts are not data-driven in any fundamental sense. State and local officials with decades spent under the sway of familiar systems and a focus on compliance with state and federal mandates constitute significant obstacles to more fundamental change. It is only in the past five or six years that many superintendents, central office staff, principals, and teachers have even embraced the principle that pupil achievement data should be a defining element of the school culture. Principal preparation continues to devote scant attention to data-related questions.¹⁶ Due to career paths in which educators have little opportunity to see how management is practiced beyond the world of K-12 schooling, there is often limited familiarity with how data might be collected or employed more aggressively. This helps foster a strong bias for data that measure “inside-the-classroom” metrics—like test results and teacher practices—rather than other dimensions of organizational performance.

Fourth, districts have done a poor job of developing and rewarding the behaviors and skills required to collect, analyze, and report information. Even when potentially useful data exist, there has to be internal capacity to examine, use, and probe them. Few districts have any spare capacity of this type. While a small team of skilled analysts could help a school district dramatically improve its operations by putting appropriate metrics into place and identifying operational inefficiencies, such analysts tend not to have a natural client base outside of the superintendent—who has many other considerations to balance. Meanwhile, such analysts are likely to have ready-made opponents among those whose inefficiencies are exposed. Thus it should be little surprise that such analysis tends to make little headway.

Finally, the current focus on “data-driven decision making,” because it concentrates on pupil achievement and school performance, has districts and schools starting at what may be the most difficult entry point. Reaching reliable inferences about what drives student achievement can be difficult even in the best of circumstances (e.g., in the case of controlled, randomized field trials). Tackling this challenge with imperfect data, under conditions fraught with potential bias and measurement error, and in a politicized environment, poses daunting challenges. While districts are busy seeking to isolate “best practices,” they are neglecting low-hanging fruit in the operational areas. In areas such as human resources, data management, and professional development there is a wealth of experience from organizations outside education that could be used to help measure, monitor, and benchmark performance. Ironically, it is by focusing on these areas of operational concern that one could most readily demonstrate the power of data to drive decisions.

¹⁵ “ERP Survey Results Point to Need For Higher Implementation Success.” Robbins-Gioia, January 28, 2002. http://www.robbsingioia.com/news_events/012802_erp.aspx

¹⁶ Hess, Frederick M. and Andrew Kelly. “Learning to lead: What gets taught in principal preparation programs.” *Teachers College Record* 109(1), 2007.

In the end, no student of government or of K-12 schooling will be surprised that political pressures can trump or overwhelm fact-based decision making. In fact, on issues like teacher performance, efficiency of maintenance operations, or school system procurement, there are sometimes involved constituencies that simply do not want certain kinds of information gathered or made public. There are no easy answers to such challenges—indeed, collecting and using data will ultimately prove as much a political challenge as a technical one. On that count, one heartening example is the success that some other public sector enterprises have enjoyed employing operational performance data. In cases such as the U.S. Postal Service, the policing examples described above, or the litany of other efforts famously flagged in the Gore Report or David Osborne and Ted Gaebler’s *Reinventing Government*, public pressure, persistence, and a commitment to rewarding reform-minded leaders has led to substantial progress even in the face of entrenched constituencies and balky bureaucracies.

WHAT TO DO?

The foregoing list of obstacles suggests just how difficult it will be for even those states, districts, charter school systems, and schools that have already embraced student testing to make the leap required to become truly data-driven organizations. Core changes on this front need to occur at the system level (whether that is a district or a charter management organization is immaterial) because what we are talking about is *management* data. That is not primarily a challenge for federal officials or state bureaucracies, except as agents to encourage, facilitate, and support system efforts. In practice, collecting and analyzing these data require a role for the state in providing funding and promoting comparability—but the primary purpose is to provide real information in real time to address real management challenges in schools and districts problems. To our minds, there are at least five takeaways for educators, reformers, and policymakers who believe in the importance of doing so.

1. Create opportunities and change the incentives. As discussed above, a crucial problem is that there is little incentive for school systems to collect the data essential for transformative management. The first step in convincing educational leaders to embrace data-based management is to allow them to actually manage. This means unwinding the webs of input-based policies and regulations governing staffing formulas, class size, service delivery, procurement, and so forth, and permitting systems to devise and deploy their own ways of doing business. To do this, state legislatures, state boards, and school boards need to find new ways to evaluate systems—monitoring district and school leaders on the basis of outcomes, employee morale, operational efficiency, and progress on these counts, and linking these measures to evaluation and compensation.

2. Get started. Much of the data needed to measure and manage performance is collected already. It may not come in convenient, automated reports, and the data sources may not “talk” to each other, but the data are there waiting to be assembled by a skillful analyst. The key to the Compstat model was not a new information technology (IT) system, but the decision to use extant crime data to guide management and the practice of holding police captains accountable for improving results. This model has in fact been extended to other city departments (CityStat) and, more recently, to school systems (SchoolStat).

Implementing such “stat” processes can happen right now (in fact, it is happening in places including Baltimore; Washington, DC; Paterson, New Jersey¹⁷; Jackson, MS; and Chicago).

¹⁷ Hu, Winnie. “Statistics Pinpoint Problems in Paterson Schools.” *New York Times*, December 2, 2007.

What is most needed is not a new computer system but talent, a focus on outcomes, and political will and organizational skill. In fact, if districts revamp their IT systems prior to implementing performance management processes, district leaders will almost certainly not get the numbers they need, but the numbers the information technology staff think they need. The only way for district leaders to truly understand what they need is by focusing on performance, identifying key processes and tasks, and then working with their teams to find smart ways to monitor those on a regular basis. Absent such leadership, it is unrealistic to expect a new IT system to fix a broken human system.

One example of the potential to get started now can be found in the Baltimore School System.¹⁸ The district started a SchoolStat process around teacher recruitment starting in the 2005-06 school year. The process did not require a fancy new data system and showed real results. Each week, starting in April 2005, the SchoolStat team met with the director of human resources to review the number of vacancies in all of Baltimore's schools by subject area, grade level, and teacher qualifications. This consistent focus on outcomes provided a needed shock to teacher recruitment teams. When they did better, it was easy to see. When they fell behind, management was able to apply pressure and assistance. Weekly meetings revealed both strategies that worked and those that did not. In August, the team tried advertising in a local community paper, and it worked. They tried holding smaller, simpler recruitment events, rather than huge expos, and that worked. They tried blast emails, and they did not work, so these were soon ended.

Baltimore reached a record low of 35 vacancies around Thanksgiving 2005. Shortly thereafter, however, vacancies shot upward. More than 40 teachers suddenly left the system after Christmas. After conducting interviews and analyzing data from three previous years, human resources discovered through SchoolStat that departing teachers were mostly recent hires who were exhausted and frustrated. This "time of disenchantment" had shown up in each of the previous three years; it had just escaped notice.

Going into the 2006-07 recruitment season, the district sought to both reduce vacancies and address the holiday "time of disenchantment." Instead of hiring to fill a projected number of August vacancies, human resources geared its recruitment goal to accommodate the "time of disenchantment" and intentionally over-hired by several dozen. These new hires were placed alongside experienced teachers who would help guide them through the difficult first year. When the holiday exodus came, the district was ready, and the rise in vacancies after the holiday season was quickly addressed. These performance improvements were accomplished without adding human resources staff or dollars, and with nothing fancier than Microsoft Excel.

While the SchoolStat model is still formally employed in Baltimore, it has evolved and many of the Stat processes have been turned over to the individual departments involved. Whether this means that the "Stat" way of thinking has been fully internalized or that the constructive pressure brought by a central "Stat" office has been removed is hard for an outsider to judge. In the end, the long-term success of such efforts is ultimately precarious and subject to the political winds, until they are firmly embedded in the culture of the organization.

3. Got money? Got talent? District leaders looking to assemble the appropriate data for active performance management face two immediate challenges. Collecting and connecting the existing data is a labor-intensive process. Then it is necessary to find those with the training and skill

¹⁸ This example was provided by Bryan Richardson, the founder of Baltimore's SchoolStat system in a personal communication.

to provide useful analysis. While these investments should ultimately save districts money, any serious move towards performance management will require more research and analytic capacity. Such spending can be politically unpopular and even runs afoul of popular calls for driving more dollars down to individual classrooms.

Private foundations are well-positioned to help districts with this challenge. They are positioned to provide start-up funding, and technical assistance for setting up performance management processes, and identify talent from non-traditional pools that will allow districts to get performance management off the ground. One promising source of candidates, for instance, is the Broad Foundation's resident fellows program, which recruits graduates from top business and policy schools, looking for proven business leadership experience. This program is much more likely to identify and recruit candidates with the skills to construct and supervise performance management systems. Funders can also help bring in external assistance to help systems develop essential tools and master critical skills. Of course, once performance management systems have shown their ability to save money and improve performance, states and systems will need to plan on shouldering the costs. Indeed, one major concern with foundations or consultants playing too large a role is that the new approach may be seen as a short-term maneuver and not be institutionalized or woven into the system's core routines.

4. State and federal governments have roles to play, too. States drive districts' core operational, financial, and student reporting requirements. If states design these requirements to capture financial data in a managerially useful way, then districts can more readily compare their costs and performance in areas from transportation to reading instruction. States can facilitate this process by creating a forum for interested school systems to meet regularly, share metrics, compare data, and benchmark their processes and results against same-state districts that are administering the same assessment and collecting data using the same state definitions and protocols. California, with its detailed standard account code structure, has taken sensible steps in this direction. On the other hand, if states design their required reporting primarily around compliance and accounting conveniences for the state, districts will adapt their internal systems to meet these requirements and opportunities to drive systematic management reform across multiple districts will be lost.

States can also reorient their own data collection responsibilities from compliance to performance management. Like local school districts, state education agencies collect mountains of data for reporting. Too often this transfer of data is a one way street. States should not only collect data on highly qualified teachers, services to students with special needs, and so forth, but feed these data back to school districts with comparisons to other districts. One small example of feeding useful data back to districts can be found in a report from the Los Angeles County Office of Education (LACOE). The 80 school districts in Los Angeles County are required by law to report their financial performance, revenues, and expenses every year to LACOE. Rather than sitting on this information, LACOE produces a report that has all of the districts' revenue and spending data per student, allowing any district that desires to benchmark itself against its neighbors on these measures.¹⁹

States could also help smaller districts capture economies of scale in the collection and analysis of data. Much of the above discussion has focused on large, urban districts that could redirect some of their central office spending to better management practices. However, smaller districts have

¹⁹ *2006-07 Annual Financial Report*. Los Angeles County Office of Education, 2007.

fewer options. They are unlikely to be able to afford either sophisticated information technology systems or high-quality business analysts. States, however, can set up these systems and provide analytic support in a much more cost-effective manner.

There is also an important role for Uncle Sam to play. For one thing, federal leadership can provide the bully pulpit and political cover to enable a few leading districts to make the case for aggressive new data efforts. By recognizing those districts that act through gestures or publications, the feds can make it easier for superintendents to act.

We would caution, though, that there is at least one role that state and federal governments should not seek to play. They should not hold districts accountable for improving their performance on the management measures discussed in this chapter. Performance metrics are tools that districts can and should use to improve student achievement and cost effectiveness. Asking state or federal officials to get involved in monitoring specific intermediate outcomes, much less attaching consequences to performance on them, implies a uniformity to improvement strategies which would limit the ability of districts to respond to their specific, varied circumstances. For instance, if the state were to reward and penalize districts on the basis of reducing turnaround time in hiring teachers, districts with an oversupply of quality teachers (for whom turnaround time was not an issue) might be forced to divert resources from more relevant challenges in response to the mandate. States and the feds should focus on ensuring that school systems are producing results—the object of interest to policymakers, parents, and voters—while leaving the granular data on system activity to the local officials and practitioners best equipped to interpret and make use of them.

5. Support management change. Finally, advocacy groups, business leaders, local media, mayors, and even governors can give district managers the political cover and support they need to move forward on performance management. Business leaders ought to not just highlight areas where school operations might be improved and support processes that allow for a reallocation of resources that can address the problems, and highlight the gains that are made in these areas as they occur.

Outside advocacy groups can help the public draw connections between seemingly non-academic management issues and student achievement. One compelling example has been The New Teacher Project's work on district hiring mentioned previously. In analyzing data from several major urban districts, TNTP has shown how seemingly mundane back-office processes can have dramatic effects on who ends up in the classroom. In 2003, TNTP examined hiring in several urban districts and found that balky human resources practices prompted roughly half of teacher applicants, especially the most qualified, to accept jobs elsewhere—with the majority of those citing delayed timelines as the rationale.²⁰ The attention prompted several districts to rethink their procedures and practices. The New Teacher Project's work provides compelling examples of how outside attention to important managerial metrics fostered awareness of an overlooked problem and changed district behavior. When local management lacks the ability, grit, or know-how to launch such efforts, external reformers can play a crucial role.

²⁰ Levin, Jessica and Meredith Quinn. *Missed Opportunities: How We Keep High-Quality Teachers Out of Urban Classrooms*. The New Teacher Project, 2003. <http://www.tntp.org/docs/reportfinal9-12.pdf>

FINAL THOUGHTS

Some might wonder whether it is realistic to expect superintendents and other school district leaders to embrace the sustained management project of collecting more and better data and using them to manage and measure performance. Indeed, some might note that one of the authors (Hess) has argued that superintendents have historically had incentives to favor fast and furious change rather than slow, incremental reforms.²¹

However, one key reason superintendents historically enacted one short-lived reform after another was an environment in which it was hard to measure outcomes and because time was short. As a result, it was paramount to *appear* to be “doing something.” This pressure can be alleviated if superintendents are accountable for measurable improvements in the near term. While it may be hard to credit a small percentile point increase in test scores to a superintendent (especially when the tests change—as they do so often), it is much easier to track improvement in teacher hiring, retention of employees, delivery of texts and materials, or data availability and utilization. By demonstrating progress in attracting quality educators, addressing needs, and wringing out inefficiencies, superintendents can win time to bring longer-term strategies to fruition.

There is a range of promising developments underway across the land. District leaders in places like New York City and Washington, D.C. have made collecting and employing operational data a priority. Charter school systems, like Edison Schools and KIPP, have taken important steps in collecting data that districts previously overlooked. Meanwhile, collaborative efforts like the Schools Interoperability Framework and vendors like SchoolNet, Wireless Generation, and S&P have brought a new level of sophistication and technical acumen to collecting data for management rather than reporting purposes. Nevertheless, if schooling is to enter an era in which data are truly a tool of breakthrough management, the real work lies ahead.

²¹ Hess, Frederick M. *Spinning Wheels: The Politics of Urban School Reform*. Washington, D.C.: Brookings Institution Press, 1999.

Frederick Hess, American Enterprise Institute
Jon Fullerton, Harvard Graduate School of Education

Center for Education Policy Research at Harvard University
50 Church St., 4th Floor
Cambridge, MA 02138
617-496-9837