CASE STUDY

Designing and Implementing a District Accountability System

by John J-H Kim, Kyla Wilkes, and Daniel Goldberg

> *School accountability is often thought to be all "stick" and no "carrot." However, our survey findings contradict this stereotype.*

Designing and Implementing a District Accountability System

| John J-H Kim, Kyla Wilkes, and Daniel Goldberg

While various systems of accountability such as testing have existed throughout the history of public education, public school districts have recently exhibited a renewed interest in designing more robust and holistic measurement systems that can drive sustained performance improvement. This focus is especially found in larger urban school districts and many state education departments that seek to respond more proactively to their base of increasingly diverse neighborhoods, parents, and students. Also fueling the interest in accountability systems is the increasing popularity of the district decentralization model as a way to drive school improvement. Many districts recognize that this model, which gives schools and school leaders greater levels of autonomy, must be complemented with a more intensive accountability system. And lastly, the 2011 announcement that the Obama administration would be granting waivers from No Child Left Behind's accountability terms has provided a compliance impetus for districts and states to begin charting their own unique "accountability paths."

Given this renewed attention to accountability, DMC recently conducted a preliminary survey of the field in order to better understand approaches to performance measurement currently being undertaken. DMC interviewed leaders of seven public school districts that have established leading accountability systems. The goal was to learn about the successes and challenges they faced and to provide insights to other districts that may be thinking about tackling similar work. Based on this survey and additional research and analysis, DMC has developed frameworks for building a performance measurement system. While our research and thinking on this topic will continue to evolve, we want to share these frameworks to provide some structure to those districts currently working on creating or refining their accountability systems.

Lessons from the Field

DMC interviewed leaders of seven mid-sized to large urban public school districts from across the country. While all seven districts have established leading accountability systems, our survey revealed that there is not a one-size-fits-all approach to accountability. Districts design systems in diverse ways based on varying goals and values. That said, consistent themes emerged that may serve as guidelines or lessons to those embarking on this type of work. While each district needs to design a system tailored to its particular needs and situation, the application of these lessons can ensure a more successful system design and implementation.

1) Position the Accountability System as a Mechanism for Driving Supports, Not Sanctions

School accountability is often thought to be all "stick" and no "carrot." However, our survey findings contradict this stereotype. All districts interviewed stressed that accountability is, first and foremost, a tool to facilitate accurate school diagnostics, to subtly direct behavior, and to provide targeted support, rather than a tool to determine school sanctions for poor performance. While all seven districts focused on accountability as a lever for improving practice, they outlined different methods for accomplishing this goal:

• Structured dialogues with school leaders: Accountability systems can provide a basis for structured and regular dialogues between principals and the central office. These dialogues may take a variety of forms, but in general, they begin in the fall with a review of performance data, peer learning, and a goal-setting process. Most districts then allow considerable flexibility for the central office to schedule subsequent conversations with principals throughout the year as needed. Given the large investment of time necessary to do this well, it is often only struggling schools that receive this form of support.

- Influencing practice through metric selection: Carefully selecting the particular metrics that roll up into the overall school quality measure can help to subtly guide the actions of school leaders. As the well-known saying goes, "You get what you measure." Many districts interviewed took this idea to heart and chose, as part of their school measurement systems, metrics that would help drive the

School accountability is often thought to be all "stick" and no "carrot." However, our survey findings contradict this stereotype.

key behavioral changes desired in their schools. For example, if a district places heavy emphasis on district-wide improvement in ELL proficiency, the district may consider heavily weighting or including a large number of ELL-related metrics in the overall "index."

- Qualitative school reviews: School quality observation reviews can augment the more quantitative school quality metrics and better highlight concrete actions a school can take to improve its overall measure on the district's accountability system. These evaluations tend to focus on "leading indicators" of student success, such as the strength of school management and a classroom culture of high expectations. For instance, a school district may use the quantitative element of an accountability system to identify a problem area and then may use the qualitative review to better understand why that problem is happening and what solutions the school or district can

implement to solve the problem. Although many of the districts interviewed expressed a desire to conduct these types of observations, time and financial expense were often cited as barriers to consistent implementation.

2) Align All Measurements of School Quality with Other District Evaluation Systems

It is critical to have a single, district-wide set of metrics to unite all actors across the system and ensure that everyone in the district is marching toward the same goals. Multiple tools or measurement systems result in unfocused and frustrated school leaders and staff. Achieving perfect organizational coherence may never be possible, but three high-impact areas are key for district-wide alignment:

- Alignment with the district's strategic plan: Alignment with the strategic plan/board accountability system was identified as essential to success. Over 70% of districts interviewed demonstrated tight coherence between the school accountability system and the district's overall strategic plan (in most cases, the board created a strategic plan with success metrics that "cascaded down" to the school level). Without this alignment, districts risk developing a fragmented system in which individual actors are confused about what goals they are working toward and how success will be evaluated. However, a perfect, one-to-one connection is not necessary or worthwhile; metrics should be added or subtracted to the school accountability system as long as they broadly reflect a district priority, as listed in the strategic plan.

- Alignment with staff evaluations: Fewer than half of districts reported that their school accountability system is directly connected with their principal or teacher evaluations, but this is an area of alignment that many are interested in pursuing. These districts believe that in order to create the appropriate incentives, principals and teachers should be rewarded based on how they drive overall improvement, as measured by the school quality measurement system.

- Alignment with legal/regulatory requirements: Many districts take precautions to ensure their district accountability system aligns with (or does not contradict) the state or national accountability systems. District measurement tools are a way to augment state or national systems by adding more holistic metrics (e.g., culture and climate, college and career readiness), but do not necessarily replace or supersede the state or national systems.

3) Engage Diverse Groups of Stakeholders in the Creation of the Accountability System

Effective communication and change management practices are among the most important aspects of a new accountability system, yet they are often overlooked. It is easy for accountability leaders to get so bogged down in technical design elements that they forget about the efforts needed to build buy-in and support for a new system. Teacher and principal buy-in is often especially important, both because these actors are closest to the students and their learning, and because they help shape public perceptions. The districts' change management efforts often focus on the following three elements:

- Involve stakeholders in the system's design: Internal stakeholder groups should be convened to provide input into the system's design. Many districts created cross-functional groups that, in some cases, included over 100 participants. In one district, the district leader sent out weekly emails keeping stakeholders apprised of the group's progress. In another, the superintendent embarked on a broad public relations tour.

- Focus on simplicity in order to increase accessibility for all stakeholders: People naturally view things they don't understand with some skepticism or even suspicion. By designing an easily calculated school ranking, districts can ensure that the system is not viewed as arbitrary or irrelevant. In one district, the emphasis on simplicity meant having to use only simple arithmetic to arrive at the overall measure; school leaders could calculate their

school's scores themselves. In another district, it meant not weighting any factor more than any other. However, this emphasis on simplicity was not at the expense of rigor or accuracy. In fact, many districts frequently screened potential metrics with regressions, dispersion analyses, and so forth, to help validate their final, simpler calculation methodology.

- Engage with stakeholders to revise the system: Creating a revision protocol early on and involving others, particularly those in the field, can increase confidence in the process. Some districts established the revision process before rolling out the accountability system in order to communicate to stakeholders that the district fully expected the system to evolve based on feedback. These protocols ranged from formal, rigorous processes to informal, ad-hoc revision policies. However, almost all shared three key features: (1) they included a broad group of stakeholders, not just the district's accountability team; (2) new metrics were almost never introduced immediately as high-stakes, but were phased in gradually as low-stakes metrics (e.g., not included in the overall rating); and (3) districts relied on complaints from the field to identify areas in need of scrutiny. After all, teachers and principals have the best sense of how particular metrics can be manipulated.

A small number of districts made the intentional decision not to offer a revision protocol, as stakeholders across these districts had expressed frustration with moving targets. This decision was the result of earnest engagement with stakeholders that created a true understanding of their needs in their particular context. As such, these specific districts focused intensely on getting it right the first time so that major ongoing revisions would not be necessary. ➔
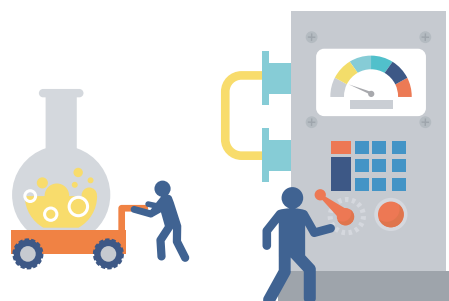
Designing a District Accountability System

How many metrics are too many? What data can be feasibly used? What types of analysis should be conducted before rollout? These are the types of questions districts grapple with when designing a new performance measurement system. DMC has been developing a framework to help districts get started. While a full toolkit for designing a district accountability system is still in progress, DMC would like to share this framework to help districts organize their thoughts and chart a clear process going forward.

The School Quality Measurement Framework

The School Quality Measurement Framework (Exhibit 1) can guide a district's thinking about potential measures and can create a much-needed common vocabulary for discussing, comparing, analyzing, and planning an accountability system for the district.

This framework begins by identifying the thematic areas or "domains" of school quality to be measured. Then, within each domain, a series of desired goals or "outcomes" can be established. These outcomes are aspirational and describe the ideal that each school should strive for. The outcomes lead to indicators, which are broad areas of measurement that can assess a school's progress toward the respective outcome. And finally, each indicator can then be broken down into concrete units of measurement or metrics. The metrics layer is highly specific and details the data the district will need to collect, track, and analyze to ultimately improve performance.

Exhibit 2 provides an example of how this framework can be used.

EXHIBIT 1: **DMC'S SCHOOL QUALITY MEASUREMENT FRAMEWORK**
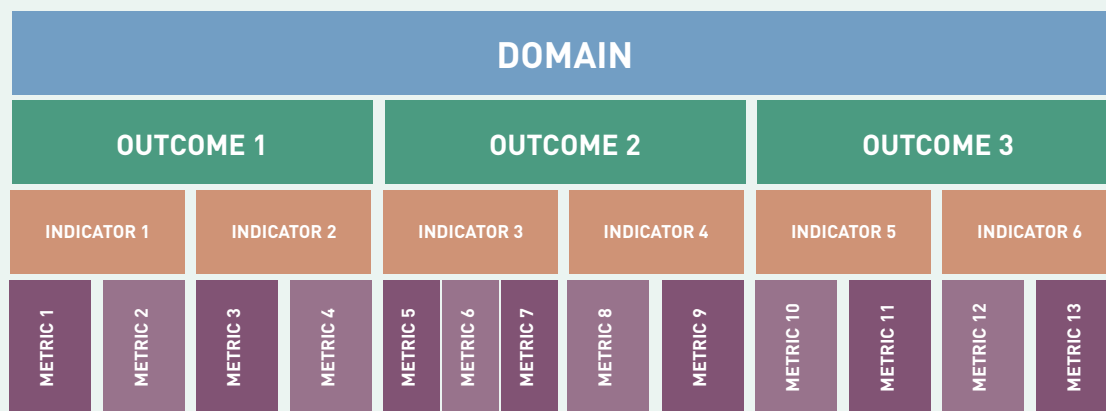
DOMAIN: Broad category or theme
e.g. Student outcomes

OUTCOME: The desired result the district and community are looking for
e.g. Academic learning gains for all students

INDICATOR: Area of measurement for progress against outcome
e.g. Student progress on summative test scores

METRIC: Data collection mechanism
e.g. School's median Student Growth Percentile on state tests

SOURCE: DMC

For example, for the domain of School Culture & Climate, we might develop the framework's cascading elements as follows:

DOMAIN	**SCHOOL CULTURE & CLIMATE**			
OUTCOME	All students are engaged and passionate about their coursework and see relevant connections to their own lives.		All students and families are safe on school grounds.	
INDICATOR	Success in college	Level of student engagement	Healthy school environment	
METRIC	Percentage of students who feel school prepared them for professional life (based on climate survey data)	Percentage of students enrolled in college the first year after high school who return for a second year	Number of suspension/expulsion incidents per 100 students	Percentage of students who report feeling safe on school grounds

SOURCE: DMC

The Four Domains for Measuring School Quality

On the surface, district accountability systems vary tremendously. Each has its own specific goals, different taxonomies, and unique groupings that are tailored to its particular context. However, in examining these more closely, the variations are more a reflection of a lack of common language than a lack of a common approach. Something of a consensus emerges that districts may use as a starting point for discussion and around which they can organize their systems. While some districts will have additional domains, these four domains are commonly found in the accountability systems of many districts (Exhibit 3):

• Student Outcomes: Measures academic achievement and growth, 21st-century skill development, social/emotional development, and college readiness, across student subgroups

• Culture and Climate: Measures whether a school provides a positive, supportive, and rigorous learning environment for all students

• Teaching and Learning: Measures instructional experience, knowledge, and quality

• Management and Organization: Measures the focus and strength of school strategy, leadership, and management supports ➡

STUDENT OUTCOMES
Indicators that measure academic achievement and growth, 21st-century skill development, social/emotional development, and college readiness (across student subgroups)

CULTURE & CLIMATE
Indicators that measure whether a school provides a positive, supportive, and rigorous learning environment for all students

TEACHING & LEARNING
Indicators that measure instructional experience, knowledge, and quality

MANAGEMENT & ORGANIZATION
Indicators that measure the focus and strength of school strategy, leadership, and management supports

SOURCE: DMC

Once the domains, outcomes, and indicators are developed, districts are faced with the difficult task of selecting metrics. This is where the rubber meets the road—while the domains and outcomes tend to be aspirational, the metrics are the concrete elements that will actually make

> **Effective communication and change management practices are among the most important aspects of a new accountability system, yet they are often overlooked.**

up the heart of the system. Given this, it is critical to put careful thought and deliberation into metric selection. In our work with an urban school district, DMC started by creating a library encompassing hundreds of possible metrics; we assembled metrics already used by the district for various reports as well as exemplars from other systems.

Once a district develops its metrics library, it will need to pare down the list in a structured, rational manner. DMC's 3R rubric focuses on reliability, replicability, and relevance, and can help guide districts in this difficult task (Exhibit 4).
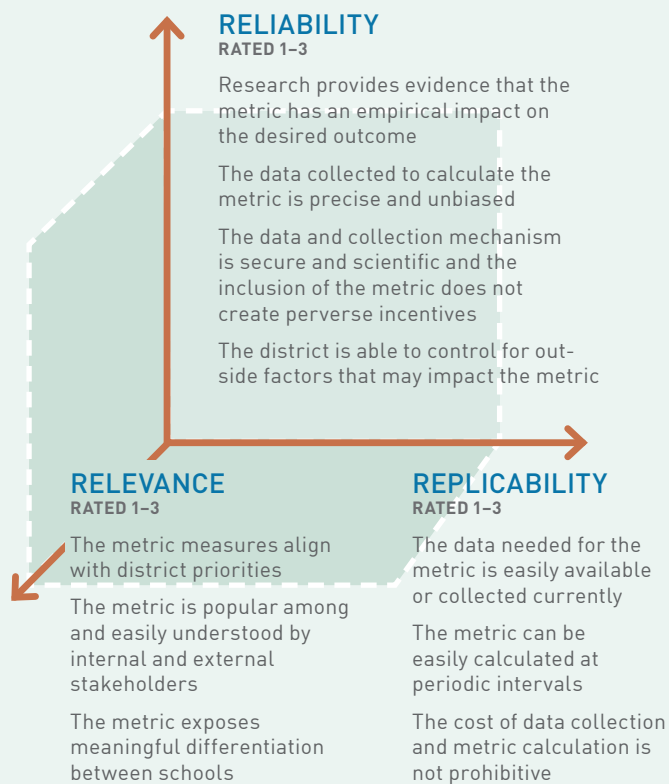
- Reliability: Does research provide empirical evidence that the metric closely correlates with the desired outcome? Is the data to calculate the metric precise and unbiased? Is the mechanism for collecting the data secure, scientific, and not apt to create perverse incentives?

- Replicability: Is the data that is needed currently collected or readily available? Can it be evaluated easily at periodic intervals? Is the cost to obtain the data prohibitive?

- Relevance: Does the metric align with district priorities? Is it easily understood by internal and external stakeholders? Does it expose meaningful differentiation between schools?

Quantitative Metric Analysis

Once metrics have been narrowed down via the 3R rubric, it is then important to run rigorous performance analyses on these shortlisted metrics to further test their validity. Two waves of analysis can be structured to garner the most insight into the strengths and weakness of any given metric.

- Metric Simulations: Multiple years of school data can be run through each metric to help the district understand the school distribution for each shortlisted metric. This analysis will provide a preliminary "sniff test" that can highlight, early on, any unexpected results or point to potentially weak data. The distribution analyses will also demonstrate whether the metric exposes meaningful differentiation between schools.

- Correlation Analyses: Sometimes metrics are shortlisted because they are expected to be proxies for performance. These assumptions must be tested using statistical tools such as regression. Only this type of technical work can confirm that each metric closely correlates to the desired outcome.

RELIABILITY
RATED 1–3

Research provides evidence that the metric has an empirical impact on the desired outcome

The data collected to calculate the metric is precise and unbiased

The data and collection mechanism is secure and scientific and the inclusion of the metric does not create perverse incentives

The district is able to control for outside factors that may impact the metric

RELEVANCE
RATED 1–3

The metric measures align with district priorities

The metric is popular among and easily understood by internal and external stakeholders

The metric exposes meaningful differentiation between schools

REPLICABILITY
RATED 1–3

The data needed for the metric is easily available or collected currently

The metric can be easily calculated at periodic intervals

The cost of data collection and metric calculation is not prohibitive

Each proposed metric can be rated on a scale of 1 to 3 on each of the three dimensions using the following scale:

1. The metric does not adhere to any of the tenets

2. The metric either
   (a) moderately adheres to all tenets or
   (b) adheres very strongly to some tenets and moderately or not at all to others

3. The metric strongly adheres to all tenets of the component

Scoring and ranking each of the proposed metrics in this manner can help the district narrow the list of metrics to be considered.

SOURCE: DMC

Next Steps

School measurement and accountability systems have the potential to drive significant change in a district. They can provide insight on how to most effectively allocate resources across schools; they can provide school leaders with invaluable data to drive critical improvement in their classrooms; and they can provide district leaders a mechanism for more clearly understanding the health of their organization. As stated previously, when a district embarks on the journey of building or reforming their accountability system, there is not one out-of-the-box solution that works at all times and for all districts. Rather, the design and implementation of a system will vary based on the specific needs of the district and the unique political context in which it operates. However, best practices do exist, and have led to success in the seven districts surveyed here. DMC hopes that by sharing the results of the preliminary survey, we can disseminate some best practices and lessons from the field.

We hope that the School Quality Measurement Framework and the 3R Metric Evaluation Rubric we have created will provide some structure and guidelines as districts are thinking through the development of accountability systems. We will continue to share our thoughts and approaches to creating accountability systems and look forward to hearing your thoughts and experiences in this endeavor.