



**OECD ACTIVITY ON IMPROVING SCHOOL LEADERSHIP**

**International Conference**

***International perspectives on School Leadership for Systemic Improvement***

**6 July 2006**

**LEADERSHIP AS THE PRACTICE OF IMPROVEMENT**

**PRELIMINARY DRAFT**

Richard F. Elmore

Graduate School of Education

Harvard University

June 2006

## LEADERSHIP AS THE PRACTICE OF IMPROVEMENT<sup>1</sup>

It is an ordinary day at Hamilton Elementary School, in an urban setting in the northeastern U.S. Students and teachers are working with a satisfying and orderly hum. The hallways are neat and clean and are hung with student work. Classrooms are busy. Students move through the hallways in a (mostly) quiet way. The school's mission statement—"Learning for All"—is posted prominently in the front hallway. The principal moves from classroom to classroom, greeted by children and teachers. Hamilton is classified under No Child Left Behind as a school in need of improvement. Its mathematics scores on the state test are significantly below the level required to meet the annual yearly progress standard; its reading and writing scores have moved in and out of that zone over the past four years. It is a high poverty school in a largely immigrant neighborhood. There are four primary languages in the school, Haitian Creole, Spanish, Chinese, and English; about one-third of the students either are, or have been in the last two years, in English immersion classes.

The problem the school is focusing on at the moment is the implementation of a new math curriculum. The system in which Hamilton resides adopted a very challenging, high level curriculum two years ago, to accompany its high level literacy curriculum. The teachers at Hamilton are having difficulty with the curriculum. It requires teachers to teach in a very different way than they are accustomed to. A typical lesson starts with a brief set-up of a problem, then students are asked to work individually and in groups to propose solutions to the problem, during which the teacher is supposed to coach students without providing direct instruction. The individual and group work is followed by some students presenting their work and others critiquing it.

The principal's observations reveal that in most classrooms teachers are unable to move away from direct instruction as the main pedagogical technique, that they focus on factual and procedural details at the expense of the math concepts, they frequently misunderstand the math concepts they are expected to teach, and that they do not expect students to be able to work at the level the curriculum requires. "There is a mismatch between the ability level of our students and the level of the curriculum," one teacher says. Just down the hall from this teacher, however, is a classroom in which the curriculum is being implemented with a great deal of skill with children similar to those in other classrooms. This teacher—a standout at Hamilton—has become a model teacher for the district; district curriculum staff regularly bring visitors to observe. Other teachers at Hamilton, however, have not observed this teacher.

The principal at Hamilton is stymied. "I've done just about everything I know how to do to engage teachers in this work, and we're just not making progress beyond the initial burst of enthusiasm that teachers felt when we first start working with higher level content. I think we're all pretty demoralized. The testing system is relentless, whether you know what you're doing or not. If we miss annual yearly progress again, we'll be in corrective action, which carries even stricter

---

<sup>1</sup> This paper was prepared for the International Conference on Perspectives on Leadership for Systemic Improvement, sponsored by the Organization for Economic Cooperation and Development (OECD), July 6, London, UK.

sanctions. I don't see how that is going to help us. We're still the same school with the same teachers and the same kids. How does it help to beat us up more than before?"

Hamilton's story is acted out in various versions across a wide variety of schools and school systems in the U.S. School administrators accept the terms of accountability for performance, even though they may argue with the specific way in which the accountability system treats them. Teachers are working more or less at the limit of their knowledge and pedagogical skill. School systems are making changes that they think will move them to higher levels of performance, often, as is the case with Hamilton's district, taking serious risks by adopting very high level curricula that require teachers and students to work in very different ways. But in many instances—probably most—they find themselves stuck at some point, not having a clear idea of what to do next.

This paper is about the role of leadership in the improvement of schools like Hamilton. Hamilton is a composite of dozens of schools I have been in over the past four or five years, schools that are struggling to do what they are expected to do under the terms of an accountability system they understand in only the most basic way. People in these schools show up for work committed to doing a good job. They are attached to the children they teach. They are, for the most part, very aware that they are not doing a good job, under the terms of the accountability system, but they don't have a clear idea of what to do differently. Often they are challenged to teach at higher levels—as the with introduction of the new literacy and math curricula in Hamilton's district. They feel challenged by these curricula, but they often—indeed usually—do not feel supported in their attempts to learn how to teach in different ways. They are, for the most part, not persuaded by the teacher down the hallway who seems to be doing better. They live in their own world with their own students. Leaders with good intentions are trying to change these schools, often themselves with only the simplest ideas of how this work is done.

This paper explores the relationship between accountability and leadership. The argument of the paper is as follows: Accountability systems work to the degree that they engage the knowledge, skill, and commitment of people who work in schools. Indeed, the success of accountability policy depends on the development of what I will call practices of improvement—explicit strategies for developing and deploying knowledge and skill in classrooms and schools. The politics of accountability tend to lead to an underinvestment in knowledge and skill, and an overinvestment in testing and regulatory control. Correcting this distortion requires changing the relationship between policy and practice, particularly around the definition and development of leadership. This paper develops a theory of accountability that is different from the prevailing view expressed in policy. It develops a model of school leadership practice consistent with this new theory of accountability. And it provides an initial working model of how school improvement works as a developmental process and how policies might be used to increase leadership capacity for school improvement. Accountability policy will not increase school performance without a substantial investment in human capital aimed at developing the practice of school improvement in a diverse population of school leaders.

**(If) Accountability is the Policy. (Then) Improvement is the Practice.**

The idea of accountability for performance has a firm grip on education policy in virtually every industrialized democracy. The social, economic, and political roots of these policies is worthy of its

own extensive analysis.<sup>2</sup> Suffice to say for our purposes that these roots run deep, and the general direction of these policies is relatively immune to change.

Policy speaks of “holding schools accountable for results.” Through the application of some combination of performance standards, assessments, classification schemes, oversight and sanctions schools, and the people in them, will come to understand what policymakers (and presumably the broader society) expects of them and, over time, change their behavior, individually and collectively, to meet these expectations. In this way, then, policy produces performance. What is interesting about this formulation of accountability policy is that, while it dominates policy discourse, there is little in the history of social interventions, or in the practice of schooling, to support it.

The first problem lies in the attribution of the effect (performance) to the cause (accountability policy). One can, of course, measure the aggregate “effects” of accountability policy on school performance—and policy researchers do this whenever and wherever they can find the data (proof of Abraham Kaplan’s first law of instruments—“give a child a hammer and suddenly everything in the world needs pounding.”) These studies are good for provoking debate and for generating and allocating political credit and blame, essential parts of the institutional behavior of democracies. But they have little or nothing to do with the actual cause and effect relationships that determine school effectiveness or performance.

We have known explicitly for at least thirty years, and probably implicitly for a good deal longer, that it is not the policy, or the program, that directly produces the effect. We have known that if policies produce any effects at all, they do so by altering the distribution of effects around some mean, typically in marginal ways. Hence, the main effect of any policy is practically meaningless as scientific construct; the distribution of effects is far more meaningful-- an idea that is hard to express in political terms. We continue, for example, to talk about the main effects of vouchers and charter schools despite the fact that these effects are typically small and variable from one study to the next, and despite the fact that most of the information about the effects of these interventions lies in their distribution, not in the main effect. We have also known that the main effect of any intervention is typically quite small, relative to the ambient noise in the larger context. It is now virtually a given that variability in effects among sites within a given intervention exceeds variability between the interventions themselves, or between the intervention and the control condition. In plain language, this means that context dominates treatment in any large scale social intervention. In the language of old fashioned analysis of variance, interaction effects dominate main effects. The effects most worth knowing about in policy analysis, and the least analyzed, are interaction effects.<sup>3</sup>

Despite these robust and repetitive patterns in policy research, policy discourse continues to focus on main effects, as if the world were organized neatly around clearly delineated policies, and as if everything important that happens in the world were directly traceable to some policy decision made by someone whose electoral fortunes depend on its success. This misconception is driven not by an

---

<sup>2</sup> See Richard Elmore, “The Politics of Accountability,” in process, 2006; Paul Manna, School’s In: Federalism and the National Education Agenda (Washington, D.C.: Georgetown University Press, 2006); Elizabeth Debray, Politics, Ideology, and Education: Federal Education Policy During the Clinton and Bush Administrations (New York: Teachers College Press, 2006).

<sup>3</sup> See Richard Elmore, “Getting to Scale with Good Educational Practice,” in Elmore, School Reform from the Insider Out: Policy, Practice and Performance (Cambridge: Harvard Education Press, 2004); Cynthia Coburn, “Rethinking Scale: Moving Beyond Numbers to Deep and Lasting Change,” Educational Researcher, 32(6), 3-12; Amanda Datnow, et al., School Reform Through a Highly-Specified Curriculum: Implementation and Effects of the Core Knowledge Sequence,” American Educational Research Journal, 36, 543-597.

understanding of the actual world in which policies operate, but rather by the incentive structure within which policymakers operate. You don't get political credit for interaction effects.

In the context of accountability and school leadership, the "main effects" view of policy has produced a number of very costly misconceptions. Not the least of these is that school performance will increase to the degree that schools and school systems "implement" accountability policy. In this view, the federal government holds states accountable, governors and legislatures hold their state agencies accountable, state agencies hold school systems accountable, school systems hold the schools accountable, and school leaders hold students and teachers accountable. We have known for a very long time, of course, that this "fidelity" or "compliance" model of policy implementation does not work, and never has. Policies "work" not to the degree that they force everyone—in this case schools—to do what the policy requires, but rather to the degree that they move different parts of the distribution of schools in a similar direction using different types of policy instruments. Again, this view seems too complex, and too nuanced for a world in which political credit is generated by doing something visible and claiming credit for it.

An even more costly misconception of the main effects view is that schools will do better if they are given clear information about their performance. In this view, delivering clear information to schools and their communities about their performance will have a galvanizing effect on the people who work in them, and will cause them to do something they would not otherwise have done to improve teaching and student performance. When I talk to my students and to groups of practitioners about this view of accountability, I ask them to imagine schools, on a grand scale, in which teachers have systematically squirreled away in their classroom closets all their best and most powerful instructional ideas and practices, saving them for the day when the accountability system whops them upside the head. Then, magically, the good ideas come out of the closet and school performance, just as magically, increases. In fact, people in schools are working pretty reliably at the limit of their existing knowledge and skill. Giving them information about the effects of their practice, other things being equal, does not improve their practice. Giving them information in the presence of new knowledge and skill, under the right conditions, *might* result in the improvement of their practice, which *might*, in turn, result in increased student performance. In the 1970s, Thomas Schelling, the Nobel Laureate Economist, called this distinction, "doing the right thing versus knowing the right thing to do." Accountability policy, as it's presently constituted, makes no such distinction.

I am constantly amazed, as I work with a variety of schools and school systems around issues of accountability and school improvement, at how little they seem to know about things that I consider to be central to the process of school improvement. I want to stress that these are schools that have been operating in a performance-based accountability system for at least a decade. They were subject to strong state accountability systems before the advent of No Child Left Behind. For the most part, they have gotten the message that accountability for student performance is their present and future. They can tell you in rough terms where they lie in some distribution of schools and districts, and they can tell you whether they are facing sanctions, and what kind, under accountability policies. In this sense, they have internalized the main message of accountability policy. But they have almost no knowledge of how to respond to accountability policy effectively—at either the school or the system level. The distribution of schools I work with is extremely bi-modal. The majority have only the smallest, most rudimentary understanding what to do in response to accountability policy. A significant minority have relatively well-worked-out strategies, and a smaller minority within this group have strategies that appear to be working. A small fraction are somewhere between the vast majority who don't know what they are doing, and the significant minority who seem to. If accountability policy were "working," in the implementation view, this distribution would look very different. If policymakers were interested in the effects of accountability policy, they would know something about this distribution, and they would be trying to do something about it. In point of fact, accountability policy

does not work when it doesn't take account of the knowledge and skill requirements for its success. These requirements vary considerably from setting to another.

What the present conception of accountability lacks is a *practice* of school improvement to go with the *policy* of accountability. Accountability policy is, for the most part, resting on a weak, unreliable and mushy base of knowledge, skill, and practice. The state of knowledge is evident in the distribution of effects, but this distribution is not part of routine discussions of the policy. In Schelling's language, it matters a great deal less in these conditions whether people want to do the right thing—for the most part, they do—but in vast numbers they don't know what the right thing is to do, or how to do it. Furthermore, and more distressingly, accountability policy itself is based on the premise that they don't need to know, because doing the right thing is all that is necessary.

In an institutional structure in which the governance of schools is increasingly defined by accountability for performance, *leadership is the practice of improvement*—like it or not.<sup>4</sup> We can talk about broader, more philosophically-grounded definitions of school leadership, but the necessary condition for success of school leaders in the future will be their capacity to improve the quality of instructional practice. In the near term, this work will have to be done in an environment which does not acknowledge the value or necessity of practice.

In sum, then, accountability policy won't work without a corresponding practice of school improvement. Furthermore, the practice has to work at a scale and to be distributed in a way that markedly alters the distribution of quality and performance among classrooms and schools. To my knowledge, no one has taken on this problem in the education system of any industrialized country. Is it worth doing, or at least trying to do? What would such a practice look like? What are the knowledge and skill requirements of such a practice? And what kind of institutional infrastructure would be required to develop and support it?

### **An Alternative View of Accountability and Leadership**

In order to get at the knowledge and skill requirements of leadership for improvement, we have to have an alternative working theory of accountability. In our work on accountability, we have found that it is useful to think about accountability as a problem of *institutional response*, rather than implementation, compliance or fidelity. Schools don't suddenly "get" accountability as a consequence of a policy being formed at some remote place and implemented at another; schools already *have* accountability. All schools, regardless of their type, status or institutional basis, have a solution to the accountability problem embedded in their existing organizational context and culture. They have answers to the questions of *to whom* they are accountable, *for what*, and *how*. These answers may not be consistent with what policymakers think they should be, but schools nonetheless have them. Some schools solve the problem by focusing on a particular group of parents, some by trying to please the superintendent, some by focusing on internal constituencies like a particular group of favored teachers. Accountability policy, in other words, doesn't "introduce" the idea of accountability to schools. It rather operates by reshaping existing modes of accountability around an alternative idea of accountability for performance to a specific, often remote, governmental authority.<sup>5</sup>

---

<sup>4</sup> Cf Michael Fullan, "The New Work of Leaders," Leadership and Sustainability (Thousand Oaks, CA: Corwin, 2005), 45-52.

<sup>5</sup> See Abelmann, Elmore, et al., "When Accountability Knocks, Will Anyone Answer?" In Elmore, School Reform from the Inside Out, 133-199; and Elmore, "How Accountability Works," in process.

So the effect of an external accountability policy depends not on whether, or how well, schools or school systems “implement” that policy, but on how they *respond* to the incentives the policy puts in place in their external environment. Accountability policies are only one among many possible signals that schools and school systems have to respond to in their environment. Furthermore, most schools operate in multiple accountability systems simultaneously and have to choose which ones to favor in any given instance. The most obvious case of this phenomenon is high schools, which have to operate in a performance-based accountability system, which judges their performance based on test scores, at the same time they operate in an attainment-based system, which judges them on the basis of their success in placing students in post-secondary institutions. These two systems are not aligned, and, in fact, are in certain critical ways in conflict with each other. Instead of asking whether schools and systems “implement” accountability policies, we ask what their responses are to the panoply of incentives they face, what the determinants of these responses are, and how they adjust to alterations of these incentives over time.

The first, most obvious, finding from our research is that schools and school systems respond differently, depending on their *capacity* and their *internal accountability*. In simple terms, which I will elaborate later, capacity is the fund of skill and knowledge that the organization can bring to bear in responding to external pressure, and internal accountability is the degree of coherence in the organization around norms, values, expectations, and processes for getting the work done. We speak of “high capacity” organizations as those that have, or have access to, knowledge and skill that can be put to use in responding to external pressures, and “low capacity” organizations as those that do not.<sup>6</sup> We speak of organizations with high internal accountability as those with high agreement around values and an organizational scheme that makes that agreement evident in practice. We speak of organizations with low internal accountability as those with weak agreement and atomized, highly variable practice. Not surprisingly, in our studies, most schools lie at the low capacity, low internal accountability end of the distribution. Perhaps a little more surprisingly, we do not find major systematic differences among different types of schools (public, private, religious, charter, etc.) on these dimensions. In our work, it matters far more what your level of capacity and internal accountability is than what type of school you are.

A school’s response to an accountability policy depends heavily on the conditions in its environment. Schools are more likely to develop capacity for high level instruction and internal accountability if they are in an institutional environment that provides support for these factors. It is not, however, a foregone conclusion that a school that exists in a supportive environment responds to that environment in ways that improve its capacity and internal accountability. Many schools get “stuck” at a given level in ways that are difficult to understand, and they seem unable to make productive use of the resources in their environment.

Think now about leadership in this view of accountability. The first thing that comes to mind is that leadership provides a focusing function, sorting out signals in the ambient environment, valuing some over others, and modeling the organization’s solution to the accountability problem around those signals. The second is that leadership is both a marker for capacity and a factor in determining the organization’s ability to mobilize and use capacity in its environment. A knowledgeable leader counts as a measure of capacity, but she also heavily influences how well the organization uses its internal capacity and develops its capacity with external resources. Likewise, leadership is both a marker for determining how internally coherent the organization is and for developing internal accountability. Knowledgeable and skillful leaders are generally (but not always)

---

<sup>6</sup> See, David Cohen and Deborah Lowenburg Ball, “Instruction, Capacity, and Improvement,” Report, Consortium for Policy Research in Education (CPRE), Philadelphia, PA, 1999; David Cohen, Susan Moffitt, and Simona Goldin, “Policy and Practice,” CPRE, January 2006.

a proxy for high internal accountability, and leadership is instrumental in developing internal accountability.

In our work on accountability and school improvement, “accountable” leaders are not passive or reactive leaders. That is, they don’t do what they are told to do. They don’t even spend much time trying to figure out and “game” the accountability systems they operate in. They operate in a more strategic frame. They use the accountability system to position themselves and their organizations in a favorable place to gain resources and capacity, and they tend to use capacity as an instrument for developing organizational coherence. Accountable leaders know that success in a performance-based incentive system does not stem from compliance, but rather from the strategic use of resources to improve performance which in turn allows them to build capacity.

Notice that I have said nothing about whether performance, as defined by the accountability system, is worth achieving in some objective sense, or against some normative principle. We have found, against conventional wisdom, that schools with high capacity and high internal accountability seem to do well on whatever the tests are, regardless of whether their instructional philosophy is aligned with the tests. It is also possible that doing well in a performance-based accountability system does not correspond to doing good. Saying that we know some of the practices of highly effective leaders under conditions of performance-based accountability systems is definitely not the same thing as saying that what they are doing is good, either for their students or for society as a whole. Nor is it clear that policymakers have any particular advantage, other than their formal institutional position, for saying what is good. For the most part, they do not have the expertise to make judgment about what is good practice educationally. Accountability policy sets a framework of incentives within which skillful leaders learn to operate; whether what they are doing is worth doing is a separate question that is argued out in the political arena and is not self-evident at any particular time.<sup>7</sup>

### **Understanding Practices of Improvement**

Leadership *practice* is what connects policy to performance in schools. It is important to understand what this proposition means in order to grasp what effective practice looks like. First, practice is *not* a personal attribute or characteristic of leaders; it is a collection of patterned actions, based on a body of knowledge, skill, and habits of mind that can be objectively defined, taught, and learned. Americans, in particular, tend to have essentialist or attribute theories of leadership: Skilled leadership, in this view, is a personal attribute, unique to the individual, like a particular posture, facial expression, or conversational style. (Americans also have essentialist views of teaching—a topic we will address later—and essentialist views of intelligence in general.) The problem with essentialist views of leadership is that they can never be generalized to scale. By definition, only a fraction of the population of potential or actual leaders have the attributes identified with effective leadership, and that fraction never equals anything like the number required for system-wide improvement. To be sure, most effective practices of leadership are initially the product of gifted individuals operating in creative ways. But these practices have to be separated from the individuals who created them in order to be useful at scale. As a gifted practitioner once said to me, “It’s only genius the first time you do it. After that, it belongs to everyone.” In order to become a practice, patterns of behavior must be objectified and separated from the individuals who use them.

Second, practice must be based on a *theory of action*, a causal model that can be falsified with evidence of its consequences. A theory of action is a set of logically connected statements that, in our case, connect the actions of leaders with their consequences for quality and performance in the

---

<sup>7</sup> The extended form of this all-to-brief argument is in Elmore, “The Politics of Accountability.”

organization.<sup>8</sup> The statements might have to do with resource allocation, with the design and structure of professional development, with the creation of collegial networks, etc. Theories of action are essential to the separation of the practice from the person. They have to be stated in order to be shared, and they have to be evaluated against evidence of their success in order to be judged. All theories of action are, of course, contingent—their actual effectiveness varies according to the settings in which they are enacted. Contingency does not, however, mean that all practices are situational, nor does it mean that theories can't be adjusted to meet predictable contingencies. For example, in order for an organization to work effectively in a performance-based accountability system, leadership has to exercise control over resource allocation, targeting time and money on developing knowledge and skill. It's unlikely that this type of theory will vary a great deal across settings, but it is likely that the conditions under which one gains control of resource allocation and the options available for the use of resources would vary considerably.

Third, practice is embedded in the particular incentive structures and particular institutional settings in which it is used. Another way of saying this is that practices cannot be generalized, except in superficial ways, beyond the institutional settings in which they are developed. Powerful knowledge of practice does not transfer, for example, from the private sector to the management of schools without considerable work in specifying, developing, and adapting it. Again, using the resource allocation example, there may be powerful resource allocation models from outside the education sector that have utility in helping school leaders think about how to manage money and time. These models do not become “practice,” though, until they are adapted to school settings, and to the particularities of incentives in those settings, and then worked into the repertoire of school leaders. Most great ideas about organization and management don't make this transition because their advocates don't have the patience or the insight to understand how practice develops.

Finally, we are interested not in practice in general but in practices that lead to school improvement. For this purpose, I will use a simple definition of improvement: ***Improvement is increases in quality and performance over time.*** (See Exhibit 1, attached) Graphically, this would be displayed as performance and quality on the vertical, time on the horizontal, and improvement would be a more or less steady movement in a northeasterly direction. This definition, of course, begs the question of what we define as quality and performance—a question I will return to later. And it raises the question of why put both performance and quality on the vertical, since, in the cosmology of performance-based accountability, performance is a proxy for quality. Again, I will address this issue later. For the time being, then, improvement is moving the herd roughly northeast. This definition might apply to teachers and classrooms within schools, to schools within systems, or to schools and systems within state or national jurisdictions.

What do I mean, then, by “***practices of improvement***”? I mean theories of action that lead to systematic increases in quality and performance over time. What I would like to do, in a preliminary way, is to sketch out what a theory of action might look like in order to illustrate the broader argument I will make later about how we might think systemically about the development of leadership in the education sector. But here it is necessary to issue a stern consumer protection warning. These insights about practices of improvement are based on my own work with schools and school systems around problems of accountability and school improvement, not a systematic body of research, which, incidentally does not exist. So what follows should be view as the provisional beginning of a conversation about leadership practice and school improvement.

---

<sup>8</sup> Chris Argyris and Donald Schon, Organizational Learning: A Theory of Action Perspective. (Reading, MA: Addison-Wesley, 1978).

***Improvement occurs in predictable stages. Practices of improvement vary by stage.***

A school in which decisions around content and pedagogy are delegated to the classroom level, in which teachers have no relationships with each other around instructional practice, in which there are no discussions among teachers or administrators about evidence of student learning is a school with extremely low internal accountability. Such schools are relatively immune to external influences of any kind because they have no receptors for new knowledge and skill and no way of using it when it is present. Moving a school like this through an improvement process requires a focus on creating occasions for discussion and analysis of instructional practice, creating a demand for new knowledge and skill, managing time and money in a way that promotes occasions for learning, and opening up classroom practice to outside influences on curriculum and pedagogy.

A school with a well-developed approach to curriculum and pedagogy, routine grade-level and content-focused discussions of instructional practice, and structured occasions to discuss student performance is a school with relatively high internal accountability. Moving a school like this requires skill in using the existing internal infrastructure to develop and sustain focus and motivate teachers to tackle progressively more difficult problems of practice. The problem with such schools is that they often lose focus, or become complacent, not they lack the wherewithal for improvement.

Notice that these two schools represent different points on a continuum of internal accountability—the first, an essentially atomized organization ill-equipped to mount any sort of response to pressure for performance, or to use any external knowledge; the second a school in the process of developing a stronger internal accountability system whose problem is how to use this system to focus on increasingly challenging problems of practice.

We could imagine a number of different points on the continuum, but these two suffice to illustrate the main issue: School improvement is a developmental process.<sup>9</sup> High performance and quality are *not a state* but a point along a developmental continuum. Like most developmental processes, this one involves more or less predictable stages. Moving a school through these stages requires an understanding, first, that there is a developmental process going on, and second, what distinguishes schools at one stage of development from another. In addition, there is a value in thinking of these schools on a continuum to demonstrate that, while the practices of improvement may vary by stage, the practice of improvement in general requires mastery of practices across stages. You cannot understand how to manage a school with a well-developed internal accountability system unless you have knowledge of how such systems develop. Likewise, you cannot create an internal accountability system unless you understand what one looks like in a more or less fully developed state.

The idea that different types of leaders are appropriate to schools at different stages of development—the contingency theory of leadership—strikes me as particularly pernicious in this context. In Massachusetts currently, for example, certain groups are advocating the training of “turn around” specialists for failing schools. These turn-around specialists have been likened to the people

---

<sup>9</sup> Here, I take gentle exception to my colleague and friend, Michael Fullan and his colleagues, who have recently published a book describing improvement using the term “breakthrough,” by which I think they mean discontinuous shifts in quality and performance that change the fundamental nature of the organization. As I have said in my foreword to this book, this has not been my experience working with schools, nor do I think the data on school improvement support this conclusion, even though I think the practices they describes in their book are consistent in many ways with what I am describing here. See, Michael Fullan, Peter Hill, and Carmel Crevola, Breakthrough (Thousand Oaks, CA: Corwin, 2006).

who deal with oil well fires—cap the well, put out the fire, return the oil field to normal. There is a presupposition that this work requires a specific set of skills that are different from those of running schools. Understanding improvement as a developmental process is the antithesis of this model of leadership. If your sole purpose is to “turn around” a failing school—and then move on to the next one—you would probably make decisions that undermine its longer-term development. Indeed, most short-term turn-around strategies necessarily involve heavy use of managerial control, rather than developing the internal connective tissue of the organization necessary to respond effectively over longer term. These strategies also have a certain heroic quality that appeals to Americans’ views of leadership—“Who was that masked man?” The hero rides out of town in a cloud of dust.

More importantly, it is difficult to shape a developmental practice of improvement if practitioners specialize in schools at a single stage of development. If the process is continuous, the practice should be more or less continuous too. Practitioners who have taken schools through a range of stages of development, and thus have developed a broad range of knowledge and skill, are a precious commodity because their practice can be captured and taught to others. These people are relatively rare, and they tend to practice in isolation. Most systems have no capacity to learn from these practitioners. In most cases, systems lack the kind of detailed knowledge of practice at the school level that is required to support improvement. In all but a few of the systems I am currently working with in the U.S., the most knowledgeable and skilled leaders are treated simultaneously as messiahs and pariahs by their superiors and peers.

Briefly put, the default culture in most schools is one in which practice is atomized, school organization reinforces this atomization by minimizing occasions for collective work on common problems, hence the school lacks the basic organizational capacity to use any kind of external knowledge or skill to improve practice. These schools exist in a myriad of contexts with a myriad of specific conditions—language groups, income groups, community cohesion and mobility, etc. As schools begin to develop toward a higher degree of internal accountability, usually through deliberate work on the development of internal processes and structures that can, in turn, be used to develop common norms and expectations for instructional practice and student learning, their success depends increasingly on their capacity to identify and respond to specific problems in their context. Schools don’t improve by following a set of rules; they improve by engaging in practices that lead them to be successful with specific students in a specific context. Hence, sustained improvement depends on the development of diagnostic capacity and on the development of norms of flexibility in practice. Leaders in these settings succeed to the degree that they engage in more or less continuous learning, and model that learning for others in the organization.

*Improvement is seldom, if ever, a smooth or continuous process.*

Typically, schools, or school systems, for that matter, do not advance through the improvement process at a more or less steady rate once they have begun. Most schools describe periodic states in which they “get stuck” or “hit the wall.” Typically, these states occur after a new practice has been adopted but before it has become deeply-seated in the organization, or after a deeply-seated practice has been in place for a while and a new problem surfaces in the organization which the practice can’t address. The early-stage version is often called “the implementation dip.” Let’s say a school adopts a new literacy curriculum, or an extension of an existing one that is designed to deal with students who have serious language deficits. Incorporating a new practice into an existing one, or displacing an old one with a new one, involves a stage of learning, challenging new practices and changing a mindset about what it is possible to do. It is not unusual for performance data to go into a stall—sometimes for a year or two—and in some cases even to decrease as the new practice finds its place in the repertoire of teachers and administrators.

Sometimes when a practice has become well-seated in the organization and it has been associated with positive increases in quality and performance, it loses its capacity to produce those increases and performance goes flat. Most of the schools I work with, having gone through what they consider to be very difficult processes of increasing internal accountability and the adoption of new instructional practices, hit a plateau within a year or two. For the people who work in these schools this is often a psychologically devastating experience. It is not unusual to hear people say things like, “I thought we had it together and it seemed like we were doing so well, and then things went flat. We don’t know what to do next.” When you look at the evidence of student performance and classroom practice in these schools a number possibilities occur. First, it is often the case that some groups of students are simply not responding to the new practice. Teachers’ sense of success is fueled by the students who are responding, but at some point the lack of response from certain students becomes a drag on overall performance of the school. Second, it is often the case that teachers and principals overestimate, by a significant amount, how much their practice has actually changed. When we do more systematic classroom observations, we often discover that some of the more challenging parts of the new instructional practice are simply not present, or not present in a powerful enough form to affect student performance. Finally, it is often the case that the original intervention wasn’t challenging enough, and the school simply needs to ratchet up its expectations for practice and performance, which involves another difficult adjustment of practice and expectations.

In general, developmental processes—biological, geological, economic, political, organizational, or human—do not follow simple straight-line trajectories. A more common pattern that generalizes across a number of developmental processes is called “punctuated equilibrium.”<sup>10</sup> People or systems might move relatively rapidly, sometimes in a discontinuous way, through a particular stage of development identified as “progress.” This stage is often followed by a period of equilibrium, in which the factors that produced the previous stage of development stabilize, and the factors that produce the next stage of development are latent and unobservable, or at least can’t be measured through the same measures one would use to gauge “progress.” Equilibrium is followed, again, by a stage of disequilibrium in which a constellation of accumulated forces produces a discontinuous development.

Learning, whether it occurs in students or adults, whether it is individual or collective, is a developmental process. As such, we should not expect it to occur in a uniform, linear fashion. The practice of improvement is the management of learning for collective purposes; hence, knowledge of development is central to the practice of leadership.

---

<sup>10</sup> The term “punctuated equilibrium” originates in evolutionary biology, specifically with a famous paper authored by Niles Eldridge and Stephen Jay Gould, “Punctuated Equilibria: An Alternative to Phyletic Gradualism,” in Schopf, T.J.M., et al., Models in Paleobiology (San Francisco, CA: Cooper and Co., 1972), 82-115. Eldridge and Gould argue that major changes in speciation occur in small populations at the periphery of large, stable central populations. These “small peripheral isolates” become a “laboratory for evolutionary change,” and result in discontinuous shifts in the major population. Eldridge and Gould portray the process of evolution as a form of continuous gradualism, punctuated by discontinuous changes. For a review of the controversy surround the idea of punctuated equilibrium, see: Donald Prothro, “Punctuated Equilibrium at Twenty: A Palenotological Perspective,” Skeptic, 1 (3), 1992, 38-47. The application of the idea to development of individuals and social organizations has been more implicit than explicit, but the idea corresponds closely to much of the current literature on adult development, organizational development, economic development. For example, Albert Hirschman, Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States (Cambridge, MA: Harvard University Press, 1970).

Improvement occurs across at least three domains: the technical, the social-emotional, and the organizational.<sup>11</sup>

At least three processes are occurring simultaneously as schools get better at what they are doing: There are changes in instructional practice that occur with some consistency across classrooms that represent more powerful forms of learning for students and adults. There are changes occurring in students', teachers' and administrators' sense of efficacy as a result of changes in practice and changes in student learning. And there are changes in the structure, processes, and norms around which the work of adults and students is organized. If you imagine development as a more or less wavy line, moving from southwest to northeast, on a graph with performance and quality on the vertical and time on the horizontal, the organization is simultaneously (a) getting better at its core functions; (b) changing the way adults and students think about their role in the process of learning; and (c) increasing internal accountability by managing the organization in progressively more coherent ways.

The practice of improvement, then, occurs across these three domains; the practice of leaders requires knowledge, skill, and fluency of practice in each, and across all three. Leaders don't get to choose to be "good" at some domains and "not-so-good" at others; they have to be competent across all domains in order to be effective. What does this look like in practice? It means monitoring instructional practice more or less continuously. It means seeding the creation of organizational structures, processes, and norms that make instruction transparent, so that it can be analyzed and changed in response to feedback about its effects. It means modeling inquiry and learning as the central dimensions of practice, creating expectations that the improvement of practice is a continuous process. It means developing practices of challenge and support that help people deal with the social and emotional difficulties entailed in improvement. And it means using the basic features of the organization—structures, processes, norms, resources—as instruments for increasing the knowledge and skill of people in the organization.

The practice of improvement consists of making the familiar strange: objectifying practice, treating organizations as instruments.

One of the most difficult aspects of mastering this practice is learning to treat existing instruction and organization in an agnostic and instrumental way. Just as our theories of leadership are essentialist, so too are our theories of teaching. We identify the person with the practice. Teachers are thought to be either "good" or "bad" depending on deeply-seated personal attributes. Teachers think of themselves as more or less coterminous with their practice; they *are* what they teach. To challenge the practice is to challenge the person. This view of teaching is, among other things, profoundly unprofessional, no, *anti*-professional—imagine a physician arguing that her surgical practice is a consequence of purely personal tastes, or an airline pilot announcing that he is making his landing approach based on his personal esthetic considerations. It is also deeply anti-intellectual—good practice, in the essentialist view, depends on who you are, not what you know or what you can do. But the main problem with the essentialist view is that it effectively precludes any possibility of improvement of instruction at scale. There are never enough people with the "right" attributes to go around. It makes what is essentially a learning process into a selection process, and, in the process, makes it impossible to treat human skill and knowledge as the main instrument of improvement.

The same might be said about the standard stance of school leaders toward their organizations. Structure, process, norms, and resources are what they are, in the essentialist view; the

---

<sup>11</sup> Robert Kegan and Lisa Lahey, How the Way We Talk Can Change the Way We Work: Seven Languages for Transformation (San Francisco, CA: Jossey-Bass, 2002); Ronald Heifetz, Leadership Without Easy Answers (Cambridge, MA: Harvard University Press, 1994).

job of the leader is to manage within them, not to treat them as instruments for making things happen. To disturb the form of the organization is to surface for critical scrutiny all the treaties, implicit and explicit, and all the accommodations that have been made to individuals, their interests and their (in)competencies. To assert that structure is instrumental to learning is to assert that the collective interests of the organization supercede the individual interests of its members. Typically, low-performing, atomized schools are organizations in name only; they are the congealed residue of private interests.

We have learned to use two specific practices, adapted from other professions, to develop the capacity of leaders to objectify their own practice and to help practitioners learn to treat the organization in which they work agnostically and instrumentally. One practice involves the use of protocols to observe, analyze, and develop practice. We observe instructional practice using protocols that focus as much as possible on the visible evidence in the classroom, *not* on the personal attributes of the teacher and not on the observer's normative stance toward what is being observed. We also use protocols for discussion of these observations that focus mainly on developing a body of evidence from which we can draw inferences about learning and student performance. We try to the extent possible to focus on the evidentiary claims that people in leadership positions make to justify their practices, or their theories of action, rather than on the personal attributes of the leaders. The use of protocols depersonalizes practice; it separates the practice from the person; it objectifies the practice, and in doing so it makes the practice something that can be changed through learning and further practice.

Another technique we use is to work with practitioners on their theories of action—asking them to state in as simple terms as possible what the main causal connections are between what they influence or control and what they are trying to achieve by way of quality and performance. Again, the value of this work is that it helps practitioners to objectify their practice, to put it in terms that someone else can understand and that, if necessary, can be used to communicate with others and to teach them what the underlying line of thought and action is. It also depersonalizes the practice, so that people can feel free to treat their own most deeply-held values and beliefs as empirical propositions that can be subject to verification through evidence on the effects of what they do.

A large part of the practice of improvement for leaders is making the invisible visible. Most people in school leadership positions are more or less socialized to a relatively dysfunctional culture. Part of that socialization process is learning to take most aspects of the organization and its culture for granted and to focus on a narrow range of things that the default culture tells you you can do. Part of the process of teaching leaders to actively manage the process of improvement is to make all the implicit rules, norms, and agreements that set constraints on action explicit, subject to analysis and change.

We work with leaders on the central cultural artifacts of their organizations: the schedule, the assignment of teachers to grade levels and classes, the use of preparation time, the use of meetings, the management of time and money, the consultation and planning processes, etc. With each of these artifacts we try to get leaders to explain how they might be used to focus instructional practice, to create challenge and support for teachers and students, to create opportunities for enhancement of knowledge and skill. Rather than treating the “givens” of the organization as constraints on action, we try to create a bias toward treating them as instruments for making things happen.

As improvement progresses, leadership refracts.

Up to this point, I have, for the sake of simplicity, committed the common fallacy of confounding leadership with role. I have spoken of “leaders” and “teachers” as if they were mutually

exclusive categories, and left the impression that leadership inheres in positional authority—in the principalship, in the superintendency, etc. Now it's time to rectify this fallacy.

As the literature on communities of practice suggests, collective learning requires distributed cognition. As networks or organizations get better at what they do, specialized expertise tends to develop in multiple sites, networks of influence develop around those sites, and leadership tends to become defined less by position and more by expertise. The literature on communities of practice tends to romanticize this process, suggesting that formal organization plays little or no role in the development and distribution of knowledge and practice. In fact, there are situations in which the development of practice relies more on social networks, and less on formal organization—punk bands, modern dance and meditation, for example. But the problem of improving practice and performance around learning is not one of these situations. Improvement of educational practice at scale requires some kind of formal infrastructure, and, as improvement progresses, leadership tends to follow the demands of learning, individual and collective, rather than the demands of formal organizations.

The metaphor of refraction is useful, if not exact. The basic idea is that the qualities and direction of light change as it passes through different media—air, water, lenses, etc. The notion is that leadership takes substantially different forms in organizations in the early stages of improvement than it does in organizations that are more advanced, and the *practices* of leadership become substantially more complex and powerful as they engage the energy and commitment of people in the organization. Organizations in the early stages of improvement rely heavily on role-based definitions of leadership, and, as they cope with the early stages of improvement they tend to import practices of improvement into traditional roles. Principals take on increasing responsibility for the instructional side of the organization, and often try to convert other administrative roles in the organization into instructional support roles—assistant principals, for example. As improvement advances, it becomes clearer that role-based definitions of leadership are inadequate, both because teachers who take improvement of their own practice seriously become more expert on instructional issues than their supervisors and because the flow of work through the organization becomes too demanding and too complex to manage exclusively from the top. So the work of leadership tends to flow out into the organization.

Notice that two things are happening simultaneously as leadership refracts—people are learning how to use their own individual expertise for collective purposes, and they are also learning a new set of knowledge and skills associated with managing work across organizational boundaries. As this process advances, people in positions of formal leadership increasingly manage less in the direct mode, and more in the indirect mode. Their work is less in direct management of the conditions that influence instruction and performance and more in managing the learning and development of people whose responsibility directly affects instruction.

There are all sorts of ways in which this process can go awry. The most common way that the process goes awry is that positional leaders underestimate the level of complexity and the demands of the work as the organization begins to improve, and they don't adapt the structure of the organization to the tasks it has to perform. As teachers start to differentiate from one another in terms of their expertise, leaders fail to acknowledge what is going on and continue to manage as if everyone in the organization were equal. It is not unusual, for example, for teachers who are highly enthusiastic and committed to professional development in the early stages of improvement to become increasingly frustrated as the process advances because the level of the professional development and the work doesn't increase as their expertise increases. Professional development becomes a compliance task rather than a learning opportunity. The organization is not flexible enough to adapt to changes in their expertise, or to use that expertise as a resource in the organization. Often people with positional

authority are threatened by the idea that others might know more than they do about key functions of the organization.

Another common way the process of differentiation goes awry is that leaders change the form of the organization without changing the nature of the work. Principals often can't give up direct management of instruction once they have had the experience of being successful at it. They deliberately design the organization so that people who have expertise are treated as specialists and not as leaders—they narrow responsibilities and treat people as subordinates rather than as active agents of improvement. Hence, they lose most of the advantages of expertise, and they narrow the range of learning available to others in the organization.

The more advanced the improvement process, the more complex the work, the more complex the processes of leading the work, and the more distributed the work becomes. The idea of distributed leadership has gained a good deal of visibility, for good reason. At its core, distributed leadership describes the way expertise and influence is distributed in schools and school systems. In general, consistent with our work on internal accountability, the more dense the networks in a given school, the more likely that school is to function effectively in the face of external demands. What the distributed leadership literature does not deal with directly, although it does implicitly, is the practice of leadership in a distributed system and how it develops over time. My sense is, from observing improving schools, is that they don't just *distribute* leadership—that is, put more influence in the hands of people with expertise—they also *develop* leadership. That is, they actively create a common body of knowledge and skill associated with leadership practice and they put people in the way of opportunities to learn it.

One common question we ask during school visits is “who’s chairing the meeting?” We’re frequently asked to observe planning meetings of various kinds in schools. One good unobtrusive measure of how leadership is defined, distributed, and developed is how the organization decides who is going to chair the meeting. Often, in advanced organizations, the principal is a participant but not the chair. When the explanations about who chairs have to do with positional authority, it is clear that there is a largely role-based definition of leadership. When the explanations have to do with who has the expertise and who’s turn it is to try out a new set of skills, then it is clear that the organization has a more developmental view.

Performance and quality are imperfect proxies for each other. Improvement requires attention to both.

Throughout this discussion I have used performance and quality more or less synonymously as indices of improvement. It is now time to unpack these concepts. Quality is a matter of professional judgment. Performance is a matter of external measurement. Both are central to large-scale improvement. An illustration and some analogies will help to illustrate.

There is a body of knowledge now about the acquisition of reading skills in the early grades. It consists of a set of more or less well-defined practices, accompanied by a body of evidence on the effects of those practices. When knowledge in a field reaches this stage of development, it becomes professionally irresponsible not to use it. The practices constitute indices of quality—that is, we can say that we expect to see certain practices in schools and classrooms as an index of quality in those settings. Evidence on the effects of the practice constitutes performance, and the external measurement of performance constitutes the core of accountability. In the healthcare sector, there are standards of practice embedded in the practice of physicians, and in the organization of practice—these standards of care define quality. We also measure the effects of practice and monitor the

performance of health care organizations as a way of making decisions about performance and cost-effectiveness.

The problem comes when, as inevitably happens, there are disconnects between quality and performance. As noted earlier, practices associated with successful reading instruction often don't work in certain contexts for certain children, and it becomes the responsibility of practitioners to figure why they are not working and to do something about it. In this case, the school may have met the quality standard for reading instruction, but is not meeting a performance standard because the quality standard doesn't cover the particular situation it finds itself in.

Likewise, many schools look much better on performance measures than they do upon inspection of their practice. Schools in general, and high-performing schools in particular, produce a large part of their performance with social capital, not with instruction. Families and communities bear a large part of the role of educating young people—often directly, such as in the purchase of tutoring services to compensate for the shortcomings of instructional practice in schools; often indirectly, such as creating pressure for attainment and performance in the lives of children independently of what the school does. For this reason, it is wise to treat the population of high performing schools with some skepticism as a source for “successful” practices, or “high quality” instruction. My experience—however limited—working with high performing schools is that it is an extremely bi-modal population. Some high-performing schools actually do contribute significantly to students' learning and performance through instructional practice. Many high-performing schools are stunningly mediocre in their practice, and produce most of what they do with social capital.

In the work of improvement, performance measures are hardly ever completely adequate for judging how well a school is doing or making decisions about how to focus resources for the improvement of instruction. Most performance measures are late in coming, and apply to cohorts of students who are no longer present in the grade where they were tested—fourth graders are tested in the spring, the results come in the fall (if we're lucky); the fourth graders are now fifth graders, and we are put in the situation of inferring what would work for this year's fourth graders from data on last year's fourth graders. In college courses on inference, we would not stand for this kind of sloppy reasoning; in accountability systems we regard it as good practice.

In addition, the grain size of external measurements is way too large to use as a basis for detailed instructional decisions. Item-level scores on tests are notoriously unreliable as a basis for making predictions about future performance. Item-level scores associated with individual students are even worse. The utility of external measurements to school-level practitioners lies in their description of aggregate-level effects, not in the fine-grained data necessary for instructional decisions.

External measures of school performance are mainly useful to higher-level authorities in making aggregate judgments about the performance of schools and school systems. Policymakers argue that the test data should be useful to schools in making detailed decisions about instruction (a) because they lack the knowledge of basic educational measurement to know the difference; and (b) because they need to justify the cost and frequency of testing by saying tests are broadly useful for decisions of all kinds.

In this situation, the rational thing for school practitioners to do is to focus on formative assessment data that is very close to the instructional process—teacher-made assessments—and to monitor the quality of instruction against some external standard of practice very carefully. This is, in fact, what most “real” professions do. They develop relatively clear standards of practice to guide their detailed daily decisions and monitor the consistency of their practice with those standards relatively carefully. Over time, they and others engage in empirical research designed to push out the

boundaries of knowledge and practice, using aggregate evidence of effectiveness to reinforce or critique existing practice.

Leadership in this situation consists of creating and sustaining the structures and processes necessary to monitor and evaluate practice with the organization against quality standards that are, to the extent possible, based on defensible criteria outside the organization. Being skeptical about the utility of external performance measures is an important part of this process. Teaching people in the organization how to manage against performance measures while, at the same time, sustaining a commitment to quality is an essential part of the practice.

In our non-degree professional development programs at Harvard University, I have taken to routinely asking the assembled administrators and teachers how many of them have taken a basic course on educational measurement. In an audience of 50 to 100 participants, the usual count is two or three. These people are usually “ringers”—they are typically assistant superintendents for measurement and evaluation. That is, they run the testing operation in their school systems. Now, imagine what the state of health care would be if practicing physicians didn’t know how to read EKGs, EEGs or chest x-rays, didn’t know how to interpret a basic blood analyses, or didn’t know anything about the test-retest reliability of these simple diagnostic measures. Imagine what it would be like if your basic family practitioner in a health maintenance organization didn’t know how to interpret a piece of current medical research questioning the validity of the standard test for colo-rectal cancer. Imagine what it would be like to be a practitioner in a health care organization in which every piece of evidence required for patient care came from a standard test of morbidity and mortality administered once a year in the organization. The organization you are imagining is a school system.

At the system level, we have created an accountability regime without creating the professional infrastructure necessary to make it work at the delivery level. This regime invites practitioners to engage in unprofessional and incompetent behavior—usually without knowing it—in the interest of doing what policymakers—who are equally uninformed—want them to do, thereby producing electoral credit. It is difficult, but not impossible to lead in such a regime. At the very least, the regime is not designed to promote the kind of leadership required to make it work.

Improvements in performance usually lag improvements in quality.

Developmental processes, of which school improvement is only one, are characteristically see-saw relationships among key variables. One variable has to advance before another one can; the latter variable advances, while the former stays constant or declines, and finally acts as a constraint on the latter. In biophysical systems increasing the food supply increases the population of a food-consuming animal, and this population out-consumes the food supply, putting a constraint on the population, which in turn increases the food supply, etc.

*Such is the likely relationship between quality and performance in schools. Our judgments of quality are relative to a particular time and a particular state of knowledge and skill. A good part of what we know about quality depends on what we learn from measures of performance. I have noticed in many schools I am in that substantial improvements in the quality of instructional practice precede by a considerable amount of time improvements in measured performance. I have also noticed that practitioners’ sense of what it is possible to do is highly sensitive to what they take to be the effects of their practice. So, for example, with the introduction of a substantially new mathematics curriculum that requires much more complex pedagogy on the part of teachers and a much more active role on the part of students, we often see in our observations that there are significant changes in instructional practice, but the formative and external measures of performance stay the same, or, in some cases decline. The problem is not that the teachers and students have gotten it wrong; they’re actually*

*working very hard to get it right. The problem seems to be that the effects of instruction on teachers' practice and student performance are more complex than a simple input-output model. The model is something more like a critical mass function—that is, the practice has to reach a certain level to displace the earlier, less effective practice, and student learning has to reach a certain level in order to displace the students' prior constructions of mathematical knowledge. Once this happens—notably after what seems like a good deal of teaching at a relatively high level—then performance seems to respond. This relationship might not exist in laboratory conditions, where the situation is more tightly constructed. Schools are social organizations and the individual performance of members is, in part, contingent on the performance of others.*

*At any rate, the implications of this relationship between quality and performance are significant for leadership practice. First, leaders have to know enough about the practice itself to know what the cognitive and emotional obstacles are to acquiring it and doing it fluently. Second, leaders have to have some systematic understanding of the various ways developmental processes work in order to be supportive and helpful to people who are struggling. Third, leaders have to be patient and to expect a possible see-saw relationship between quality and performance, watching for evidence of changes in student learning before they become evident in external measures of performance.*

It is also important to observe that the design of most accountability systems is no friend of the developmental view of leadership practice. The incentive structure of most accountability systems puts a premium on direct and immediate effects on performance. This view is based on exactly no knowledge of how the improvement process actually works; there is no empirical basis for the performance targets in accountability policies because there is no research relating performance to the presence of other factors in the environment of schools. The gap between what good leadership practice might look like and what the accountability environment signals it should look like is, at the moment, quite wide.

## **PRINCIPLES OF LEADERSHIP DEVELOPMENT**

Accountability works to the extent that it is supported by practices of improvement. Performance is a collective good. Its value exceeds that which can be produced by any single individual, organization or system acting in its own self-interest, narrowly construed. For reasons too complex to develop here (but developed at length in other places), the politics of accountability leads predictably to an under-investment in the capacities required to produce the collective good called performance. There are deep systemic reasons why we have tended to underinvest in the very capacities required to make accountability systems work. Correcting this situation requires changing the relationship between policy and practice, particularly around the definition and development of leadership. The following four principles are designed to provide some initial guidance in how leadership might be defined more clearly as a collective good and made more productive in a regime of performance based accountability.

### ***(1) Honor the principle of reciprocity.***

Fundamental to the political economy of accountability is the principle of reciprocity. Accountability is essentially a contractual relationship in which a principal contracts with an agent to

act in a particular way—in this case, to produce a certain level of performance. In order to work, this relationship has to be beneficial to both parties—the principal receives the benefit of the agent’s performance (in the case of policymakers this benefit accrues largely in the form of electoral credit) and the agent receives both the authority that inheres in acting for the principal and whatever the material benefits are in the transaction. As we have noted above, accountability is a special case of the principal-agent relationship, since, unlike many such relationships, the principal cannot depend on the agent to know what to do. If practitioners knew how to solve the performance problems they faced in schools, accountability systems would be unnecessary, or at least would look very different from what they do. So the principal-agent relationship is complicated in the case of accountability systems by the fact that in order to get what the principal demands, the principal and the agent have to act cooperatively to build the knowledge and skill of the agent to do the work. This is where the principle of reciprocity comes in.

The principle of reciprocity, in its simplest form, says that *for every unit of performance I require of you, I owe you a unit of capacity to produce that result*. In practice, this means that accountability for performance requires investments in capacity that are equal to the expectations they carry. Now, there are lots of complexities we can introduce to the principle in its simplest form. We can assume, for example, that there is lots of unused capacity in the system that can be mobilized to produce results, the accountability system has to first exhaust that unused capacity before the simple form of the principle applies. We can assume that teachers really do know how to teach at higher levels, but for some reason, they simply aren’t doing it. We can assume that principals really do know how to manage resources in the service of improvement, but for some reason, they’re not doing it.

Another complication around the principle of reciprocity is who gets to decide what kind of capacity is needed to produce a given level of performance. If we leave the decision to policymakers, it is clear that they have very strong incentives to underinvest in capacity and to treat accountability systems simply as instruments to mobilize unused capacity. If we leave the decision to practitioners—especially to practitioners who, themselves, don’t know what they need to know in order to improve performance—it is likely that we will either overinvest or invest in the wrong things.

The solution to the problem of reciprocity in the real world of school improvement has to be incremental. First, policy itself has to acknowledge the principle of reciprocity—something accountability policies presently do only marginally, if at all. Second, practitioners and policymakers have to build a strong institutional relationship around the link between capacity and performance. In auto emissions control policy, for example, there are elaborate institutional settings and processes for arguing out what it is feasible to produce, given the existing evidence on the characteristics of the internal combustion engine and the government’s goals for pollution control. No such settings exist in education. In their absence, policymakers are more or less free to set performance standards anywhere they want, and practitioners are forced to live with the consequences. If the policy succeeds, the policymakers claim credit. If it fails, they blame the practitioners. Either way, they get electoral credit. We tend to underinvest in capacity because there is little or no discipline in the system to enforce the principle of reciprocity.

The infrastructure that would be required to enforce the principle of reciprocity would be one that combines the expertise of researchers to track the effects of various practices of improvement combined with the engagement of networks of practitioners to develop, test, and evaluate their own use of these practices. Absent a body of evidence at a level of specificity that would inform practice, it would be very difficult to say what the capacity requirements of strong leadership are and whether the government is meeting its responsibilities under the principle of reciprocity.

## ***(2) Treat leadership as a human investment enterprise.***

The model of leadership that emerges from the practice of improvement has three important characteristics: (1) It focuses on the practice of improving the quality of instruction and the performance of students; (2) It treats leadership as a distributed function rather than as a role-based activity; and (3) It requires more or less continuous investment in knowledge and skill, both because the knowledge base around instructional practice is constantly changing and because the population of actual and potential leaders is constantly depleting and replenishing itself. In this view, leadership is a knowledge-based discipline. The practices associated with leadership exist independently of the people who use them, and they are subject to constant testing against the rigors of practical work and evidence of effectiveness. Leadership does not inhere in the personal characteristics of the individual; it inheres in the knowledge, skill, and behavior of the individual.

Accountability as it is conceived thusfar requires more or less continuous improvement of performance. If accountability systems were fairer than they are now, the requirement of continuous improvement would apply equally to nominally high performing schools as to nominally low-performing schools. The model of leadership that applies to continuous improvement is one in which the system is constantly investing in the capacity of people at all levels to master and lead the improvement of instructional practice.

One thing I have noticed about education as a system, relative to knowledge-based enterprises in other parts of society (health care operations, consulting firms, law firms, research and development organizations, information technology organizations) is that education systems typically have almost no human resource management function. Human resource management in the typical school system consists of hiring new teachers and administrators. Professional development, if it exists as an administrative function, is typically located in another part of the organization. And supervision and evaluation is shoved down in the organization to the school level, where it becomes routinized and disconnected from anything having to do with instructional practice. So whereas most knowledge-based enterprises have a unified structure for recruiting, hiring, inducting, mentoring, training, supervising, and promoting individuals, all organized around the goals of the system, educational systems have a host of separate functions that typically work at cross-purposes to each other. Policy aggravates this problem by treating credentialing as regulatory issue and professional development as a grant-in-aid activity. The current requirement in No Child Left Behind that there be a “qualified teacher” in every classroom is a travesty in a system that has no capacity to manage human resources systemically.

In a knowledge-based human resource management system, recruitment into leadership positions begins the minute a novice teacher contacts the organization for employment. Every individual would be evaluated not just on their qualifications for employment, but also on their potential to assume leadership in the organization. Every novice teacher would be supervised by an experienced teacher who modeled not just excellent instructional practice but also practices of observation, analysis, problem-solving and work with peers that characterize successful practices of improvement. Every intermediate teacher would be given some leadership responsibility in some part of the instructional improvement enterprise, under the mentoring and supervision of a more expert practitioner. Teachers would be given more or less continuous feedback on their practice, not only in the instructional domain, but also in the practices of improvement— working with peers on instructional issues, taking leadership responsibilities in groups, creating and demonstrating instructional solutions to pressing problems of performance, etc. Teams of teachers and novice administrators would be given responsibility for working through organization and management functions around problems of instructional practice—designing schedules that provide time for teachers to pursue common work, designing group work around instructional practice, designing

induction activities for novice teachers, etc. From the group of more experienced teacher leaders, the potential cohort of principals would be created and these individuals would operate in a setting outside their present school under the supervision of another principal with increasing responsibility for school-level management functions. System-level administrators would be recruited from the ranks of teacher leaders and school-level administrators with strong instructional knowledge and strong managerial skills.

The incentive structure in this system is the same as in any knowledge-based enterprise—positional authority follows the contours of expertise, it is the responsibility of experts to induct, socialize, and manage novices, evaluation and supervision centers on mastery of practice, and lateral accountability is as important as vertical accountability because much of the work is accomplished in groups.

In the context of current practice in public education, this kind of human resource management system sounds extreme. In the context of most knowledge-based enterprises, it is routine. Students who take my courses, and who are changing careers from law and health care, are dumbfounded at how weak the human resource management systems are in schools and school systems. I have heard every possible explanation for why it is impossible to create such a system—it costs too much, it combines too many functions with too much specialized knowledge in a single place, it requires skills and knowledge that people in the organization don't have, and (my favorite) it undermines the positional authority of too many people. In fact, the small population of schools that advanced to the outer edges of the practice of improvement have had, of necessity, to create human resource systems that look very much at the school level like a good human resource management system would look at the system level. The problem is not that people can't invent these systems. The problem is that the broader managerial and policy environment is unresponsive to them when they are invented—yet another example of how accountability systems don't work in the absence of knowledge and skill, and how accountability systems don't fill the knowledge and skill gap in the absence of explicit attention to the principle of reciprocity.

*(1) Invest in social capital<sup>12</sup> around practices of improvement.*

*Knowledge and skill in accountability structures are collective goods, not private goods. That is, the knowledge and skill necessary to improve the performance of schools doesn't belong to those schools, or to people who work in them, it belongs to the system as a whole—if, that is, accountability is about systemic improvement. Accountability systems, to be sure, send mixed signals—on the one hand, they seem to want to induce competition among schools as a way of spurring performance, on the other they seem to regard performance as something that should be common to all schools. If knowledge and skill becomes a private good, then accountability works not to promote systemic improvement but simply to shift schools around in the distribution of performance, or to advantage one set of actors over another. Policy should be concerned, as I argued above, as much or more with shape of the distribution of performance as it is with the aggregate effect of policy on performance. If this is true, then policymakers have no choice but to treat knowledge as a public good.*

*The problem of how to create and deploy knowledge in the leadership of improvement is a classic problem of social capital. The knowledge itself doesn't reside in the individuals, it resides in the relationships among individuals engaged in the practice. What a teacher or principal “knows” has no*

---

<sup>12</sup> The most direct application of the idea of social capital to school improvement and accountability is Anthony Bryk and Barbara Schneider, Trust in Schools: A Core Resource for Improvement (New York: Russell Sage Foundation, 2002). See also, Partha Dasgupta and Ismail Serageldin, eds., Social Capital: A Multifaceted Perspective (Washington, D.C.: World Bank, 2000).

value, except insofar as it can be used to create or enhance knowledge and skill in others. One teacher's success working through a particular problem of practice has immediate value for her and her students, but it does not produce value for the school in which she teaches without intentional action on the part of her colleagues. One school's success has immediate value for the students, practitioners and parents in that school, but its public value is limited by its position as one unit in a system, and therefore its public value is limited to its direct beneficiaries. In order for an accountability system to produce performance as a public good, it has to be accompanied by a system of social relationships that take knowledge out of the private domain and make it public—within classrooms among schools, among schools, and among systems of schools within a larger polity.

The analysis of practices of improvement above suggests that the kind of skill and knowledge required to create improvement is very specific—specific to the instructional issues practitioners are trying to solve, specific to the stage of improvement in which a school finds itself, specific to the particular mix of students in the school, etc. Decisions about the adaptation and use of knowledge have to be made very close to the ground. This aspect of improvement practice suggests that investments in social capital should be densest at the level of the classroom and school, and should become less dense at higher levels of the system. Networks of teachers working with researchers and curriculum developers on the solution of particular problems of practice, networks of principals operating across schools around common problems of practice, vertical teams of administrators and teachers trying to solve problems of systemic improvement—all of these kinds of social networks exemplify what social capital formation would look like in an accountability system focused on improvement. Such networks exist. I myself have been involved in the formation of two such networks—one for school superintendents focused on systemwide instructional improvement, one for principals in a single school system also focused on systemic improvement in that system. Insofar as these networks exist, however, they tend to exist on a purely voluntary basis, with no supporting infrastructure from public authorities who are responsible for accountability. Voluntarism is good, but it also feeds variability, and variability feeds inequality of access to knowledge. Voluntarism cannot be the basis for systemic improvement.

Accountability systems tend, by the natural application of political incentives, to drift in the direction of regulation and hierarchical command and control. This drift moves policy away from investments in the social capital necessary to create, nurture and expand practices of improvement. It also tends to push leadership in the direction of positional authority and hierarchical control. Governments do what they know how to do unless otherwise disturbed. One thing governments know how to do is to promulgate regulations, run enforcement processes, and administer sanctions. This view of governmental action could not be further from the role required of government in the creation of social capital.

### **(3) Build the strategic function.<sup>13</sup>**

One thing that is striking about schools and educational systems, at least in the U.S., is the absence of anything that might be called a strategic function. I work with a number of large U.S. school systems—systems with anywhere from 50,000 to 1 million students. These are organizations that spend in the 100s of millions to billions of dollars. They employ thousands to tens of thousands of

---

<sup>13</sup> Stacey Childress, Richard Elmore, and Allen Grossman, "Promoting a Management Revolution in Education," Working Paper, Public Education Leadership Project, Harvard Business School and Harvard Graduate School of Education, 2005. This paper grows out of four years of work with senior administrators in nine large U.S. school systems. It develops a framework for the strategic function in education systems, and it provides a structure for a larger set of working papers and case studies that we have used in the project.

people. They make decisions affecting the lives of millions of people. In no system I currently work with is there anything I would call a well-developed strategic function. By this I mean a part of the organization that is dedicated solely to making the various pieces of the organization fit together coherently around a single strategic vision of what kind performance, at what level, is required for the organization to sustain itself, and what the specific organizational structures, resource allocations, and investments in human skill and knowledge are required to make those commitments work. I know of several superintendents and principals who are brilliant strategists. They personally have a vision for how the pieces of the organization fit together, and they act on those visions. If they were mowed down by a bus tomorrow their schools and systems would revert immediately to the default model that existed before they came. Furthermore, their brand of leadership typically exists of having people in the organization try to guess their intentions, and read their actions, in order to figure out what the big picture is. As a consequence no one really knows exactly what the big picture is. These leaders often create documents they call strategic plans. The rhetoric in these plans is usually inscrutable to all but the small circle of people who drafted them. The point is that, while they might be brilliant strategists, they are not brilliant leaders because they have not institutionalized the strategic function in their organizations.

I must say this absence of a strategic function is practically unique to educational organizations. Public utilities, health care organizations, research and development organizations, and, for heaven's sake, even universities now have strategies that have actual consequences for issues like who they hire, how they spend money, and how they organize themselves to do the work. I find it very peculiar that educational systems, supposedly working in a tight and unforgiving accountability environment, still have developed anything resembling a strategic function.

Strategy is a key responsibility of leadership, broadly defined. When we work with superintendents and principals on these issues, we start by having them state a theory of action by which what they do, on a daily basis, influences instructional practice. Then we work with them on how they treat the basic functions and structures of the organizations—schedules, supervision, collective work, allocation of discretionary resources, relations with community and parents, for principals; curriculum and professional development, supervision, administrative structures, accountability processes, and discretionary budget decisions for superintendents. We then try to develop practices that convert these “givens” in an organization into strategic resources that can be used to accomplish purposes, and we try to help them use the resources of their employees to help them come up with ideas about how to manage against specific performance objectives. The mindset required to do this work is counter-cultural to school people. That is, strategy requires you to see everything in the organization as instrumental to the achievement of some collective goal. Educators are, for the most part, are used to thinking of everything as “given” and then trying to find some modest source of leverage from outside to move the organization. This work is, they report to us, the most difficult work that practitioners have done.

As leaders become socialized to the work of improvement, it is important to have some place they can learn how to create and manage a strategic function at the school and system level. This does not mean creating a specialized part of the organization with “strategy” over the door. Nor does it mean that they should get better at producing school improvement plans, which are almost universally useless as strategic documents. It means that members of the organization participate in the development of a strategy that embodies real decisions about real resources, structures, and processes that have real consequences for the way the work is done. It means that people are taught to think about the time of people in the organization as money that has already been spent, and as the most important resource the organization has to improve its work. It means that people are taught to treat existing structures and process as malleable and as instruments for getting things, not as expressions of the property rights in individuals in the organization to particular benefits and

privileges. It means that everyone—including students and parents—should be able to say what the basic commitments of the organization are, how they will be accomplished, and what their role is in achieving them.

As with other practices of improvement, this one is best learned on the spot, confronting real problems, facing real people, under real conditions—not in university classrooms, and probably not even, except in an introductory way, in workshop settings. The most powerful way to learn strategy is through social networks in which the work is targeted on things that have to be done and in which there is lateral accountability and support among colleagues to do the work. Which brings me to my final recommendation.

(4) *Locate the learning as close as possible to the work.*

Elmore's second law is, "the effect of professional development on practice and performance is inverse to the square of its distance from the classroom." (Elmore's first law, for those who are interested, is, "children generally do better on tests they can read than those they can't.") As a general design principle for the development of leadership, the influence of learning on practice is greater the more direct and immediate the application to practice. In my own work, I insist, if I am to engage in longer-term professional development activities with administrators, that they commit to a regular regime of classroom observations, systematic analyses of instructional practice, and collective problem-solving around the practice of instructional improvement. In graduate coursework on instructional improvement, we spend first five weeks of a thirteen week course watching, analyzing, and drawing inferences from video tapes of teachers teaching. After this, students are required to do their observational study outside of class, and their final research project has to be a plan for improvement that includes direct work with teachers and administrators in schools. I do these things because I want to drive home the view that professionals are experts who have a practice; education is, except in rare precincts, a profession without a practice. Anyone who pretends to lead an organization whose core functions involve decisions about practice, should herself have a practice that directly connects to that core function.

For the most part, we have no working theory about how to organize learning for professionals in this field. The institutional structure of learning is largely driven either by entrenched institutional interests (cash-for-credit certification programs at colleges and universities) or by entrepreneurship that is disconnected from any strategic vision of how to generalize learning in the field (leadership academies). There is no clinical practice in the field on which to base judgments about more and less effective ways to learn it, so anybody's idea with a little money behind is as good as anyone else's. This way of organizing knowledge is typical of pre-professional occupations. They refuse to specify a norms of practice, and, for the most part, they refuse to exercise real control over entry to the occupation, and therefore they exercise only modest influence over the quality of practice generally.

My bias is to make investments in learning as bottom-heavy as possible—putting most of the resources not in formal institutions but in professional networks, anchoring the work in instructional practice rather than in managerialism, and making the criteria of success the improvement of instructional quality and student performance. Insofar as established institutions have a role to play in the development of practices of improvement, they should have to compete for the business by demonstrating that they have the expertise and the capacity to manage network-based learning systems. The first priority should be on improving the practice of people in the field, and using the knowledge gained from this effort to train others in the field for leadership positions.

*One thing that policymakers would discover if they were to underwrite such a learning system is that the costs of accountability are considerably higher and of a different order from what they presently assume. Accountability systems are currently functioning in an almost purely regulatory mode. The human investment side of the work is either ignored altogether or funded as a purely symbolic gesture. Putting money behind the development of more powerful instructional practices and more powerful practices of improvement requires a different view of what accountability is about—it is essentially a major human investment strategy—and what would constitute success—major improvements in the level and distribution of performance require major work on the practices that produce performance.*

## **LEADERSHIP, POLICY AND PRACTICE**

Accountability policy “works” by affecting the way schools, as organizations, respond to external signals about their performance. The key determinants of that response are the capacity of schools to produce high levels of instructional practice reliably, which is a function of the knowledge and skill of teachers and administrators, and the internal accountability, or coherence, of the organization around norms, expectations, and routines for getting the work done. Improvement—defined here as increases in performance and quality over time—is the process by which schools move from relatively atomized and ineffective organizations to relatively coherent and effective organizations. The process of improvement, like all developmental processes, is neither continuous nor linear; it looks more like a process of punctuated equilibrium, periods of significant increases in performance, followed by periods of consolidation. Leadership, in this context, is primarily about (a) managing the conditions under which people learn new practices; (b) creating organizations that are supportive, coherent environments for successful practice; and (c) developing the leadership skills and practices of others. Leadership of improvement, if it is to result in the improvement of quality and performance at scale, must be conceived as a practice—a collection of patterned actions, based on a body of knowledge, skill, and habits of mind that can be objectively defined, taught, and learned—rather than a set of personal attributes. As improvement advances, leadership refracts; it ceases to follow the lines of positional authority and begins to follow the distribution of knowledge and skill in the organization.

The single greatest weakness of accountability policy as it is presently constructed is its failure to invest adequately in the human knowledge and skill required to form strong practices of improvement. From a policy perspective, the agenda for developing leadership is primarily an agenda of creating the institutional structures that support the development of the knowledge and skill to lead improvement, and the social capital that connects the individuals’ knowledge and skill in ways that contribute to the development of practices of improvement. The most effective investments will be (a) close the ground—that is, in networks and institutional arrangements that connect people in classrooms and schools with the knowledge required to their work, and with other practitioners faced with similar problems of practice—and (b) will create human resource systems that develop the knowledge and skill of educators from the earliest stages of entry to the profession to the latest, rather than focusing on a single role or a single career stage.

The role of public policy in this domain has to focus on the collective goods dimensions of the improvement of practice—that is, the dimensions of the problem that cannot be addressed by individuals and schools working alone in their own spheres. More specifically, public policy has to

create the legal and institutional framework that requires the education profession to say what its practice is, to create the infrastructure by which knowledge about content and pedagogy will be made available to practitioners, and to create the career structure required to develop human talent for leadership roles. Public policy has to manifest its commitment to the principle of reciprocity by providing financial support, and constructing the infrastructure, for the improvement of practice equal to the demands for performance that accountability policy makes on individuals and institutions. And public policy has to begin to discipline its expectations to empirical evidence on what schools can achieve by way of performance, given the resource investments and the state of practice on the ground.

Exhibit 1: Improvement Processes

