

## **Complexity, accountability, and school improvement**

Jennifer A O'Day

17313 words

1 October 2002

[Harvard Educational Review](#)

293

Volume 72, Issue 3; ISSN: 0017-8055

English

Copyright (c) 2002 ProQuest Information and Learning. All rights reserved. Copyright Harvard Educational Review Fall 2002

### Complexity, Accountability, and School Improvement 1

In this article, Jennifer O'Day builds on her earlier work defining and examining the standards-based reform movement in the United States. Here, O Day explores accountability mechanisms currently associated with standards-based reform efforts that "take the school as the unit of accountability and seek to improve student learning by improving the functioning of the school organization. "She examines such accountability mechanisms using the theoretical framework of complexity theory and focuses on how information travels through complex systems, with the understanding that information, its existence and usage, is key to improving schools. Drawing on work conducted with researchers at the Consortium for Policy Research in Education (CPRE), she contrasts the Chicago Public Schools' outcomes-based bureaucratic accountability approach with the combination of administrative and professional accountability found in the Baltimore City Schools. She argues that the combination of administrative and professional accountability presents a much more promising approach for implementing lasting and meaningful school reform.

This article, like much of the conversation among reformers and policymakers today, is about accountability. Everywhere you turn - from Congress to the statehouse to local communities and parent groups - some people are trying to make other people more accountable for some thing in education. However deafening at times, these cries for accountability should not surprise us. Public education consumes over \$400 billion in public revenues. It is reasonable that the public and its representatives want to know where the money is going and what it is producing. Are educators doing what they are being paid to do? Are administrators responsible in how they are spending money? Are children engaged and learning what they need to know?

Such questions are hardly new. In the early days of the common school, for example, teachers were closely scrutinized and called to task even for such personal habits as demeanor and dress, as well as for their duties in the classroom. Meanwhile, student accountability - in the form of grades and report cards - has been around for even longer, while fiscal accountability for districts came to the fore with the rise of federal programs such as Title I in the 1960s. Yet, as many observers have noted, the current emphasis on and efforts toward educational accountability represent a departure, or evolution, from previous practice. Researchers at the Consortium for Policy Research in Education (CPRE) have labeled this evolution the "new accountability" and have analyzed its various components or manifestations.<sup>2</sup> Central among these are the emphasis on student outcomes as the measure of adult and system performance, a focus on the school as a

basic unit of accountability, public reporting of student achievement, and the attachment of consequences to performance levels (Elmore, Abelman, & Fuhrman, 1996; Fuhrman, 1999). Other analysts have delineated typologies of educational accountability (Adams & Kirst, 1999; Darling-Hammond & Ascher, 1991; O'Day & Smith, 1993; O'Reilly, 1996), noting differences among administrative/bureaucratic accountability and legal, professional, or market accountability systems with respect to who is holding whom accountable for what. In each case, reformers and observers assume that the goal of current accountability-- based interventions is (or should be) the improvement of instruction and student learning.

This article focuses on one class of the current accountability mechanisms - those that take the school as the unit of accountability and seek to improve student learning by improving the functioning of the school organization. My goal is to place the current trends and typologies of school accountability into a theoretical framework drawn from the literature on organizational learning and adaptation. I begin with a discussion of schools as complex systems, focusing on the role of information and interaction in system change and on critical mechanisms of, and barriers to, the use of information to improve schools. In particular, I argue that accountability systems will foster improvement to the extent that they generate and focus attention on information relevant to teaching and learning, motivate individuals and schools to use that information and expend effort to improve practice, build the knowledge base necessary for interpreting and applying the new information to improve practice, and allocate resources for all of the above. The most widespread and well-developed policy approach for addressing these four tasks is the outcomes-based bureaucratic model of school accountability evident in most states and districts and codified in extreme form in the recent federal legislation, the No Child Left Behind Act (NCLB) of 2001.<sup>(3)</sup> Using the Chicago experience as an example, I discuss the promise and limitations of this most pervasive approach to school accountability in light of the theoretical framework.<sup>4</sup> I then consider professional accountability as a potential alternative, arguing that the combination of administrative and professional accountability is the most promising approach for fostering organizational learning and improvement in schools. In the final section I draw out several implications of the discussion for the refinement of accountability policies.<sup>5</sup>

### School-Based Accountability: Tensions and Problems

This discussion starts from the premise that school accountability mechanisms by their very nature seek to increase student performance by improving the functioning of the school organization. Mechanisms of school accountability vary from jurisdiction to jurisdiction, but generally include the establishment of some target level of performance (aggregated across the school, though targets may include disaggregated benchmarks as well), with consequences (and sometimes assistance) meted out to the school unit for achieving or not achieving the target performance. Whatever their differences in terms of the targets or the consequences, policies that take the school as the basic unit of accountability must contend with a number of inherent problems if they are to effect organizational change. I raise three such problems here and return to them at the conclusion of this article.

Problem 1: The school is the unit of intervention, yet the individual is the unit of action. The first of these problems concerns the relationship between collective accountability and individual

action. School accountability by definition targets the school unit for monitoring, intervention, and change. But schools are collections of individuals, and to the extent that the needed change involves the behavior of the members of the organization, it must occur ultimately at the individual level. That is, individual teachers, administrators, and parents must in some way change what they are doing in the hope that this will change what students do (individually and in interaction with teachers) in such a way as to increase or deepen student learning. Schoollevel accountability approaches bank on school members' identification and interaction with their organizational environment to motivate and direct individual action. In other words, such policies assume that targeting the school unit will generate the necessary and desired changes in the behavior of individuals within that unit. This assumption leads to two questions: How will school accountability mechanisms reach beyond the collective level to mobilize such changes among individuals? What conditions need to be in place for this connection to occur?

Problem 2: External control seeks to influence internal operations. Just as individuals operate within schools, schools are nested within larger systems and environments. New accountability approaches, by their very nature, seek to influence from the outside what goes on inside schools. Moreover, such policies assume that external forces can play a determining role in changing the internal workings of schools. The limitations of such assumptions, however, have provided grist for the vast literature on policy implementation in education.<sup>6</sup>

The heart of the issue is the problematic relationship between external and internal sources of control and the implications of this relationship for organizational learning and improvement. Organizational systems have several mechanisms at their disposal to control the behavior of individuals and subunits. Two such mechanisms are formal rules and normative structures. Large systems - like public education - tend toward bureaucracy and reliance on rules. Teachers work a certain number of hours a day, teach their classes in a prescribed order, and follow a variety of district, state, and federal mandates. But rules decreed from on high often have little impact, especially when it comes to the core technology of teaching and learning (Elmore, 1996; Marion, 1999). One reason is that externally generated rules may come up against the power of an organization's internal norms of behavior. Normative structures inside schools, such as the privacy of classroom practice, are often the determining factor not only in the implementation of policy, but, more importantly, in the school's overall effectiveness in fostering student learning.<sup>7</sup> The resulting questions for school accountability policies are profound. What is the appropriate and most effective balance between external and internal control? What are the mechanisms for achieving this balance? Can external accountability measures influence the development of internal norms that are more conducive to improving student learning?

Problem 3: Information is both problematic in schools and essential to school improvement. The third problem in school accountability concerns the nature and role of information in school improvement. Indeed, information is the life-- blood of all accountability mechanisms: one accounts to someone for something, and this accounting gets done by conveying information. Current school accountability policies, such as public reporting of student test scores, assume that, armed with accurate information about the achievement of students in the school, stakeholders and participants in the instructional process will take whatever action is necessary to improve learning outcomes. But again, this simple assumption raises a host of questions, the answers to which are anything but straightforward. What are the most effective forms and uses of

information in the school improvement process? What is the potential for the external accountability system to generate and disseminate the information needed to accomplish the accountability goals? What are the motivational and learning links between information on the one hand and individual and collective action?

These three problems - collective accountability versus individual action, internal versus external sources of control, and the nature and uses of information for school improvement - undergird this paper on school accountability. To illuminate their interrelationships, I turn to theories of organizational complexity and adaptation, focusing on the central role of information in both accountability and school improvement process.

## Complex Adaptive Systems and Change: A Framework for Understanding School Accountability

### Complexity Theory and Organizations

Complex systems theory draws on parallel developments in the physical, biological, and social sciences that challenge traditional notions of linear causality and externally imposed or predetermined order.<sup>8</sup> Complexity theorists use the term complex adaptive systems (CAS) to describe "a world in which many players are all adapting to each other and where the emerging future is very hard to predict" (Axelrod & Cohen, 1999, p. xi). A few central concepts from complexity theory as applied to organizations are particularly instructive in understanding both the potential and the limitations of current approaches to school accountability.

#### -Interaction and Interdependence in Complex Adaptive Systems

A central characteristic of complex adaptive systems is the interdependence of individual and collective behavior. CAS are defined as populations of interacting "agents" (be they cells, animals, people, organizations, or other systems), each of which pursues a limited set of strategies in response to its surroundings and in pursuit of its goals. The systemic nature of a complex adaptive system derives from the patterns of interaction and to some extent from a commonality of strategies among like agents (Axelrod & Cohen, 1999; Marion, 1999).

In a school, for example, teachers interact with students, with other teachers, with administrators, with parents, and so forth. In each of these interactions, the individual actor follows his or her own goals and strategies, which differ from those of other actors to varying degrees. One teacher may seek to develop her students' mathematical reasoning while another is more focused on keeping students safe and off the streets. These teachers' differing goals may be manifested in somewhat different strategies in the classroom. Similarly, while one student may show up because he wants to hang out with friends and avoid the truancy officer, another has his eyes on an elite university. Again, students' particular activities will vary according to their goals. This variation among individual actors and their strategies is central to the notion of complexity and adaptation.

Equally central, however, are the constraints placed on that variation by the interaction of actors within the organization and between them and the larger environment. Thus, what a teacher does

in her classroom depends greatly on what her students do, and vice versa; and what individuals in one classroom do is shaped in part by their interaction with individuals in other classrooms, in the hallways, and on the grounds of the school. The more frequent and powerful those interactions are, the more influence they are likely to have on the behavior of individual actors. Conversely, the more weakly connected - or loosely coupled (Weick, 1976) - the organization, the more independent and self-determined the actions of individuals and subunits within it. The nature and strength of the patterns of interaction are thus key to understanding the relationship between individual and organizational behavior and change.

#### -Stability and Change

In complex systems, the strategies that individual agents and organizations pursue reflect both stability and change over time - with both these characteristics being manifestations of learning. As new members come in to the organization, they become socialized into (i.e., learn) its "code" - the languages, beliefs, and routines that make up the dominant behaviors of the organization and define it as an interactive system (March, 1991); hence, the stability of the system. But organizational codes and practices also change. Such change occurs through the selection, recombination, and adaptation of strategies, based on information derived from the interactions of system members with one another and with their environment. Thus, second-grade teacher Mrs. Cardenas moves Anton to a different seat because she notices that he is distracted by activity outside the window near his current seat. Meanwhile, Mr. Arthur and his third-grade colleagues introduce cross-class reading groups, based on feedback and modeling from a literacy coordinator working with the school. And Washington Elementary selects a new mathematics curriculum after receiving notice from the school board about a new testing policy. As Huber (1991) explains: "An entity learns if, through its processing of information, the range of its potential behaviors is changed. This definition holds whether the entity is a human or other animal, a group, an organization, an industry, or a society" (p. 89).

#### -Information and Learning

At the heart of the learning process in any complex system is the role of information and the movement of information among agents and subunits through patterns of interaction. Several characteristics of this information are important to note.

First, information and adaptation derive from variation. Without variation among the agents of a system or in the surrounding environment, there is little information on which to act. For example, students - individually and as a group - vary from teachers in knowledge and maturity, vary from each other, and vary from their prior behavior over time as they grow and learn. Much of teachers' activity is bound up with responding to these variations in and among students. At the system level, monitoring and responding to variation is also at the heart of current accountability efforts, as differences in reported test scores become the basis on which progress, success, and failure are judged.

Second, information in this model is not a static entity that is disseminated unchanged. Rather, information is dynamic, constantly changing as it is interpreted based on prior experience, recombined with other information and knowledge, and passed on through interaction with

others. Interpretation is "the process through which information is given meaning" (Daft & Weick, 1984, p. 294). As such, it depends on "a person's prior cognitive map (or belief structure or mental representation or frame of reference)" (Huber, 1991, pp. 102-103), which is socially constructed and varies across organizational units having different responsibilities. One implication is that the process of spreading information will itself become a source of variation - as in the classic children's game "telephone" (Marion, 1999). In essence, the message changes each time it is passed on. Spillane (2000, 2002) has found, for example, that variation in the implementation of state standards derives in part from variation in the way that district leaders interpret and then convey those standards to school personnel, who then reinterpret them based on their own context and frames of reference? A second implication is that the meaningfulness of the information generated by the system will vary in relation to the knowledge and skills of the users. To the extent that such knowledge and skills are weak or are unequally distributed, so too will be the meaning and usefulness of the accountability information.

Finally, not all information leads to learning and change. As implied above, for information to be useful, members of the system must first have access to it, through interaction with other members or the environment.<sup>10</sup> Moreover, if they are to incorporate the information into their cognitive maps or repertoire of strategies, they must attend to it and must have sufficient knowledge and stability to interpret it. Action does not necessarily follow, even once learning occurs, as this step often requires motivation and resources beyond those necessary for the learning itself. As I will discuss later, these elements - access, attention, knowledge, motivation, and resources -- are all essential. A breakdown in any one of them may disrupt the connection between information and change.

#### -Learning and Improvement

Even when learning and change occur, they do not necessarily lead to improvement. This may seem counterintuitive, given popular notions about learning and "learning organizations" (Senge, 1993), but this conclusion flows from the conception of learning and of complex systems just described. Complexity theorists Robert Axelrod and Michael Cohen (1999) use the term adaptation to indicate when learning and change lead to "improvement along some measure of success" (p. 7). This concept of adaptation is highly consistent with the goals of school accountability systems, which seek improvement in terms of specific assessments of student performance. In complex systems, however, the relationship between learning (an "increase in the range of potential behaviors)" (Huber, 1991, p. 89) and adaptation is obscured. For one thing, individual agents vary in their goals and thus in measures of success; adaptation for one might not be adaptation for others (Axelrod & Cohen, 1999).

In addition, interaction and interpretation of information make learning unreliable. Agents and organizations may misinterpret feedback from their environment, causing "superstitious learning" (Levitt & March, 1988).<sup>11</sup> If individuals and organizations act on that superstitious learning - for example, using it to select a strategy - the results may be maladaptive rather than adaptive; that is, they may lead to a decline rather than an improvement in relevant measures of success. Finally, learning at one point in an individual's or system's experience may inhibit learning and change at another point. As individuals and organizations gain competence in certain activities through learning, they may actually decrease their range of potential strategies

(Levinthal, 1991). By repeatedly selecting strategies that have led to success in the past and interacting with other entities based on those strategies, an organization reduces variation, thus potentially reducing future learning and adaptation. The end result is that individuals and mature organizations over time may get caught in "competency traps," becoming prisoners of their own past success (Levitt & March, 1988).<sup>12</sup> One might argue that the general inertia of the educational system is a reflection of such competency traps.

### Barriers to Improvement in Schools

This discussion of complex adaptive systems has implications for how we think about the potential relationship between accountability policies and school improvement and, more specifically, about the organizational barriers that accountability policies must overcome in order to foster successful adaptation. These barriers center on the generation, interpretation, and use of information for school and system improvement.

#### -Too Much or Too Little Information

Teachers and schools are constantly bombarded by information and by demands to do something about that information - requests from parents, notices from the central office or federal and state departments, publishing companies, external advocates and programs, community groups, counselors, etc. In many schools, teachers' and students' work is subject to continual interruption as others try to thrust new information upon them. What's worse, much of the information is irrelevant to the improvement of instruction and learning. It merely distracts attention and resources from what is supposed to be the main work of school personnel and students. Sifting through the morass to find that which is likely to lead to improvement requires time, resources, and knowledge that school personnel may not possess. Unable to make productive choices, some teachers and schools move chaotically from one demand or source of information to another, with insufficient focus and time to learn. The classic "Christmas tree school" is an example of this pattern.<sup>13</sup>

Alternatively, teachers and schools may metaphorically and literally close the door on new information, shutting out the noise. This is a coping strategy that potentially allows them to focus, but it also leads to isolation. Such isolation prevents their having sufficient opportunity to encounter variation and the information it engenders, and thus little opportunity for learning. Norms of private practice in education and loose coupling (Weick, 1976) throughout the system reinforce this isolation.<sup>14</sup> On the one hand, loose coupling prevents failure in one unit (school or classroom) from cascading throughout the system (Marion, 1999; Weick, 1976), but at the same time it may inhibit learning that can lead to improvement.

Consider a veteran teacher typical of those in many U.S. urban schools. Extended survival in the system implies that he or she has learned and selected strategies that have been successful by at least some measure of success - perhaps student learning, perhaps social competence, perhaps simply control. As described earlier, one such strategy has been to shut out the cacophony of information and distracting interactions by literally closing the classroom door on them. Other teachers in the school, many of them also veterans in the system, have adopted a similar approach, so that the school has become a collection of virtually private domains demarcated by

the walls surrounding each classroom. Moreover, protecting the sanctity of those domains has become an integral part of the school code. New teachers quickly learn this code if they intend to continue in the organization; they develop their own strategies for what they do in the classroom, leaving their colleagues to do the same. The structure of the school day and the school building reinforce this pattern. The teachers survive, and over time survival increasingly becomes the measure of success, not only for individuals but for the school as a whole. That this means of survival has been enhanced through isolation has implications for future change. In a sense, one could say the "egg-crate" structure of U.S. schooling, with its separate and isolated classroom structures designed to buffer each component from change, is a kind of competency trap. That is, separating teachers and students in independent and isolated classrooms is a strategy that has worked in the past for survival, control, and sorting, but that may not work well if the measure of success changes (e.g., to high levels of student learning for all students). The "trap" of the egg crate is that such isolation prevents teachers from taking advantage of the variation in their individual classroom strategies in order to learn from one another and select more successful ways of doing things.

On the one hand, external sources thrust an overload of information on schools and school personnel, while on the other hand, isolation creates a lack of information sharing among teachers. What seems to be needed is some middle ground - as in the story of Goldilocks and the just-right porridge. For Goldilocks, only the porridge of just the right temperature could ensure taste, consumption, and eventual satisfaction. Likewise, teachers, given the right amount of information, will be able to attend to it, interpret it, and use it for learning and adaptation. The question is, how will the accountability system ensure that school personnel receive an adequate amount of information, can interpret the information, and have the ability to focus on what is most appropriate for improving teaching and learning?

#### -Complexity and the Problem of Attribution in Schools

The amount of information encountered is only one part of the problem. Also of concern are the kinds of information available, how that information is interpreted, and whether the interpretations will lead to learning and to selecting strategies that lead to adaptation.

Theorists (e.g., Levitt & March, 1988) tell us that organizations are oriented toward targets and adapt based on feedback (information) about those targets. Discrepancies between observed outcomes and aspirations for those outcomes can provide motivation for change (Simon, 1986). Such is certainly the assumption of outcomes-based accountability. First of all, agents within schools vary in their definition of a target (sometimes it is just getting through the day!), and schools also have multiple and changing targets. Moving a school community from an emphasis on discipline and order to a focus on student learning, for example, is difficult and represents this sort of change. Second, even taking a single generally agreed-upon goal, such as independent reading by the third grade, does not remove the ambiguity. Other targets remain, agents have differing views about what independent reading entails, and measurement is difficult.

But perhaps most important, even in an ideal situation where the goals and measures are clear, the complexity of interaction patterns inside and outside the organization and of the learning process itself makes attribution of cause and effect difficult and unreliable. Consider the

following hypothetical example: Only 15 percent of third graders in Bryant Elementary are reading independently by time of the spring assessment. Do we conclude that the third-grade teachers are not teaching appropriately? And if so, in what particular ways is the instruction inadequate? Are the textbooks too difficult? Too easy? Too boring? Perhaps the low performance is due to constant interruptions during reading time or lack of order in the school. Maybe the problem is that the students do not see any reason for reading, or perhaps they speak a language other than English and do not have the requisite English vocabulary. Perhaps the real problem lies in first- or second-grade instruction, in the fear induced by violence in the neighborhood, or in the low expectations of the adults. Perhaps it is all or some combination of the above. Alternatively, if the majority of students at Bryant are doing well, we are not necessarily any closer to understanding the cause of their performance. Is it class background? School selection processes? Motivated teachers? Effective instruction? All of the above? Indeed, it is often more difficult to pinpoint the cause of success than it is of failure.

The difficulty of attribution, and thus of selection and prediction, is endemic to complex systems. Complexity theorists use terms like circular causality and multiple interactions to discuss the problem of attribution where individual actors are responding to (and providing) feedback from (and to) multiple sources (Marion, 1999; Weick, 1976). Mutual adaptation inevitably leads to mistakes in attribution.<sup>15</sup>

The tendency toward faulty attribution in education is commonplace. It is exacerbated by the lack of accurate, valid, and appropriately detailed information on outcomes. For example, a test given in the spring with the results reported to teachers (about their now former students) in the fall does not solve the attribution problem for teachers. That problem is also exacerbated when those involved in interpreting the results know little about the type and distribution of strategies and resources used to produce those results, or when they lack the opportunity and knowledge to reflect on and explore alternative interpretations of the information they have. As we will see, school accountability systems generally seek to address the first of these conditions - the lack of information on outcomes - but often ignore the other information needed for appropriate attribution.

#### -Faulty Incentive and Resource Allocation Structures

Adaptation in schools is also inhibited by incentive and resource structures that undermine motivation and the opportunity for organizational learning or that preclude the adoption of more productive strategies. These inhibitors include such things as incentives that pull attention and effort to goals other than student learning,<sup>16</sup> insufficient time or other resources for collaboration and sharing information about instruction, and human resource systems that reward mediocrity and concentrate the most knowledgeable teachers where they are least needed. One might think about such allocation as a "strategy" of the larger system in which the unit or individual is embedded. Thus, we would ask whether the school has an effective strategy for allocating resources to foster student learning. Does the district? As with other aspects of the system, resource allocation is dependent on accurate and reliable information. Again, the goal of the accountability system should be to supply that information.

#### School Accountability Framework

This discussion suggests a framework for analyzing the potential impact of accountability-based interventions on school improvement. School accountability mechanisms will be successful in improving the functioning of school organizations to the extent that those interventions are able to:

- \* Generate and focus attention on information relevant to teaching and learning and to changes in that information as it is continually fed back into and through the system. Note that in order to alter what happens in classrooms, this focus must occur not only at the school level, but at the level of individual teachers as well. Interaction patterns are likely to be very important in the generation and spread of such information.

- \* Motivate educators and others to attend to relevant information and to expend the effort necessary to augment or change strategies in response to this information. Central here is the problematic relationship of collective accountability and individual action. Motivation must ultimately occur at the individual level, but it is likely to be dependent in part on the normative structures of the school as well as on individual characteristics of educators and students.

- \* Develop the knowledge and skills to promote valid interpretation of information and appropriate attribution of causality at both the individual and system levels. As discussed above, learning takes place through the interpretation of information, whether that information is data from a student assessment, research on reading instruction, or observation of a colleague's lesson. Interpretation is dependent on prior learning and is constrained and informed by such. Data often remain unused because educators lack the knowledge base for interpretation and incorporation of the new information. If accountability systems are to be successful, they will need not only to build knowledge and skills for interpretation in the short run, but also to establish mechanisms for continued learning through use of information generated by the system.

- \* Allocate resources where they are most needed. Information at all levels can promote the allocation of resources - human and material - to where they are most needed. A classroom teacher might reallocate resources by spending more of her time and attention on a student she sees is having trouble understanding a new concept. Similarly, district administrators might move additional resources to a low-performing school or one taking on a new challenge. To what extent does the accountability system encourage allocation (or reallocation) of resources to foster student learning based on information generated?

### Bureaucratic Accountability and School Improvement

How do current school accountability policies fare with respect to the framework outlined above? That is, to what extent do accountability policies generate and focus attention on relevant information, motivate the use of that information for individual and collective change, contribute to the knowledge base needed for appropriate interpretation of that information, and use information to provide adequate resources where needed? To answer this question requires that we first define what is meant by "current school accountability policies." Here I turn to analyses of the new accountability in education (Elmore et al., 1996), as well as the more expansive typologies of educational accountability (Adams & Kirst, 1999; Darling-Hammond & Archer, 1991; O'Reilly, 1996).

Accountability systems, according to these and other observers, differ in large part by the way they respond to four central questions: Who is accountable? To whom are they accountable? For what are they accountable? And with what consequences? Fuhrman notes that one of the distinguishing characteristics of the new accountability in standards-based reform is that the "who" in this formulation is generally the school unit, and as previously indicated, this article focuses exclusively on school-level accountability.<sup>17</sup> In addition, while school accountability policies differ in their particulars from jurisdiction to jurisdiction, the "to whom" designation almost universally refers to the district and/or state agencies.<sup>18</sup> That is, schools as collective entities are accountable to the higher levels of the educational system. In this respect, such policies represent a form of administrative (O'Day & Smith, 1993) or bureaucratic accountability (Adams & Kirst, 1999; Darling-Hammond & Archer, 1991). However, they differ from traditional forms of bureaucratic accountability in one very important respect: they hold schools and school personnel accountable not for delivering designated educational inputs and processes but for producing specific levels or improvements in student learning outcomes. Thus, they are examples of what might best be termed "outcome-based bureaucratic accountability."<sup>19</sup> In this section, I will use Chicago as an exemplar of this form of school accountability.<sup>20</sup>

#### Addressing the Framework: Outcomes-Based Accountability, Chicago Style

The Chicago Public Schools (CPS) provide a particularly useful model to illustrate this framework, since the Illinois legislature amended the Chicago School Reform Act in 1995 to include specific provisions for school accountability. Following those provisions, the Chicago School Board designated school-level targets for student performance and instituted sanctions (probation and reconstitution) for schools falling below those targets. The district has now accumulated six years of experience with school accountability. Several colleagues and I have had the opportunity to follow the design, practice, and results of this system over the past several years. Our investigation has included in-depth interviews with business, political, education, and community stakeholders; analyses of school improvement plans and the planning process; interviews and shadowing of support providers; school case studies; and multilevel analyses of survey and achievement data on all CPS elementary schools since 1994 (two years prior to the implementation of sanctions). Our data are thus both rich and varied. They provide an inroad into understanding the links between outcomes-based school accountability and school improvement.

At first glance, CPS and similar school accountability systems seem to address well the criteria laid out in the framework above. Below I discuss how the four components of the school accountability framework are reflected in the CPS accountability system.

Attention: On the most basic level, these accountability systems call attention to information on student outcomes by designating a particular indicator (or indicators) of those outcomes and by defining specific performance targets. In Chicago's case, the focus is sharpened by the district's use of a single indicator - the Iowa Test of Basic Skills (ITBS)<sup>21</sup> - in only two subject areas, reading and mathematics. Moreover, the targeted performance benchmark is simple, measurable, and clear: schools must have at least 20 percent of their students in grades 3-8 or 9-11 reading at or above national grade-level norms in the spring administration of the ITBS, or be declared "on probation." Attention to the outcomes is further enhanced through school planning and reporting mechanisms that single out reading and math scores and require all schools to provide

information on how they will increase student performance in these areas. Such mechanisms establish priorities in the organization and thus should help school personnel sift through the usual information overload to focus on that most directly related to student achievement and improvement strategies.

**Motivation:** Chicago's policy, like outcomes-based policies in other jurisdictions, provides motivation for this sifting process and related improvement efforts by attaching consequences to the outcome targets. For all schools, these consequences come in the form of public and administrative scrutiny of reported school outcome data. For schools falling below the target, sanctions include the stigma of the "probation" label, decreased autonomy as local school councils lose authority to select their principals, additional requirements for planning, monitoring, and assistance, and potentially even reconstitution or reengineering, both of which entail involuntary changes in personnel.<sup>22</sup> Policy designers believe that even the threat of such sanctions will increase educator motivation and efforts to improve student learning.

**Knowledge Development:** Attending to outcome information is of little value if school personnel do not know how to interpret it, and motivation to act will produce nothing if educators do not know what actions they should take. Recognizing the need for site-based knowledge and skill development, CPS administrators instituted an elaborate program of assistance for schools, including mentoring for principals, help with business and school improvement plans, and professional and organizational development provided by external partners. A particularly interesting feature of the CPS design of external assistance is the district's response to the tension between internal and external sources of control discussed earlier. In an effort to balance these sources of control and enhance normative buy-in among school personnel, CPS allows probation schools to select their own partners from an approved list. The policy designers hoped that this selection process would both enhance motivation and ensure that support providers pay attention to the particular conditions in each school.

**Resource Allocation:** Finally, funding for this assistance demonstrates a major way in which the district has used information generated by the accountability system to allocate resources. Low test scores trigger the targeting of discretionary monies - initially from the district surplus and then from federal programs, including the Comprehensive School Reform Demonstration (CSR) program and class-size reduction - to probation schools.<sup>23</sup> The district covers 100 percent of the cost of the first year of assistance, 50 percent of the second year, and the school bears the full cost in subsequent years. In the first two years of the probation policy, CPS spent \$29 million for external support alone.<sup>24</sup>

Other jurisdictions, both states and other districts, have set in place similar systems - in part in response to the accountability requirements incorporated into Title I of the Elementary and Secondary Education Act (ESEA) in 1994.<sup>(25)</sup> Of course, jurisdictions vary considerably in the specifics of their policies. One important dimension of this variation is in the definition of targets. Some jurisdictions have set improvement targets for all schools and rewards for those meeting or exceeding targets. The intent is to focus and motivate improvement in both higher performing schools and lower performing ones.<sup>26</sup> Jurisdictions also vary in the forms of assistance provided and in the consequences attached to either high or low performance relative to the targets. Yet, despite these variations, the general school accountability model is the same

in that it defines certain expected levels of performance, designates schools as high or low performing based on student assessments, requires planning to focus attention and coordinate action in the school, provides assistance in some form, and administers sanctions for continued failure to improve.

#### -Impact of Outcomes-Based Accountability

Although experience with and research on school accountability are still in the early stages, some evidence of its impact is beginning to accumulate. Our data from Chicago (Finnigan & Gross, 2001) and CPRE research in Kentucky and Charlotte-Mecklenberg, North Carolina (Kelley, Odden, Milanowski, & Heneman, 2000), indicate that teachers are working harder in response to the accountability measures and are more focused on externally set student-learning goals. In addition, many systems (e.g., Boston, San Diego, Tennessee, and California) are using school-level data on student outcomes to allocate additional discretionary resources where they appear to be most needed. Some jurisdictions, such as New York City and Baltimore, have even put in place special monetary incentives to attract and retain highly skilled teachers and principals for the lowest performing schools (Westat, 2001).

There is also evidence of an impact on achievement, as measured by standardized tests. In each of the first four years after instituting its school accountability policy, the Chicago schools posted increased scores in both reading and mathematics, though reading scores have begun to level off since 2000 (Chicago Public Schools, 2002). Similarly, Kentucky, California, Texas, Tennessee, and other jurisdictions have claimed that their accountability policies have resulted in higher student achievement. However, some observers question whether increases in test scores really indicate higher levels of student learning or whether later scores have been artificially inflated by concentrated "teaching to the test" and increased familiarity with test questions and format.<sup>27</sup> In Texas and Tennessee, rising scores on the National Assessment of Educational Progress (NAEP) seem to validate similar increases on the state assessments, but in these and in all other cases it is difficult to attribute such increases to school accountability mechanisms.

Schools respond unevenly to outcomes-based accountability policies and this unevenness may be directly tied to internal conditions in schools that make them more or less able to use the information generated by the accountability systems. The CPRE research team led by Richard Elmore and Leslie Siskin, for example, has found that schools that are better positioned in terms of their socioeconomic composition (i.e., higher SES students) and their prior academic performance respond more readily and coherently to the demands of external performance-based accountability systems than those schools less well situated (DeBray, Parson, & Woodworth, 2001; Elmore, 2001). This research suggests that lower performing schools actually lose ground relative to the well-positioned schools once an external accountability system is instituted.

Our research on the lowest performing schools in Chicago extends the CPRE analysis, identifying variations in responses among schools that might all be considered less well positioned - that is, among those at similarly low socioeconomic and achievement levels.<sup>28</sup> The first indication of this variation is a rapid bifurcation in the achievement trends for all elementary schools placed on probation in 1996, despite comparable initial achievement. More specifically, one group of schools - those that came off the probation list by spring of 1998 - posts a

significantly sharper increase in scores than those schools that remained on probation after 1998. Multilevel analysis of survey data for this rapidly improving group suggests that they differed significantly from other probation schools along several dimensions of initial school capacity: peer collaboration, teacher-teacher trust, and collective responsibility for student learning (Gwynne & Easton, 2001). Referring to the earlier discussion of organizational complexity, we might surmise that the first two of these dimensions - peer collaboration and teacher-to-teacher trust - reflect stronger patterns of interaction among the organizational agents (teachers in this case). The third dimension - collective responsibility for student learning - suggests that attention and effort in these schools were already directed to a higher degree to student learning and were bolstered by the normative structure of the school - what Elmore and his colleagues call "internal accountability" (DeBray et al., 2001; Elmore, 2001).

That schools with such patterns of interaction and attention would be more successful at adaptation makes sense in light of the earlier discussion of complexity, information, and learning, as these patterns would facilitate the dissemination, selection, and interpretation of information relevant to student learning. Meanwhile, the failure of other schools to show similar improvement and their propensity to stay on probation for as long as five years or more suggests something about the limitations of bureaucratic accountability for catalyzing school improvement in low-capacity schools that is, in schools lacking the internal accountability structures or knowledge base to generate or use information well. Indeed, our qualitative data on the policy, assistance, and individual school response point to significant limitations of bureaucratic outcomes-based accountability for fostering school improvement.

#### Limitations of Bureaucratic, Outcomes-Based School Accountability

The earlier discussion of complex adaptive systems suggests that adaptation (improvement) is based on the feedback (information) agents receive from one another and from the environment, on the interpretations and dispersal of this information through patterns of interaction in the organization, and then on the invention, selection, and recombination of strategies to produce improvement along some measure of performance. Our data from Chicago indicate several ways in which bureaucratic outcomes-based accountability may inhibit - or at least fail to promote - widespread organizational adaptation.

##### -Inadequate Information

A central limitation of school accountability in Chicago and elsewhere is that the nature and quality of the information produced and dispensed by the system are simply inadequate for effective organizational adaptation and learning. Three aspects of this inadequacy stand out.

The first limitation of the information dispensed in an outcomes-based school accountability system concerns the validity of the outcome measure on which improvement is to be based. Much of the criticism of Chicago's model of school accountability has centered on the use of a norm-referenced basic skills test that is not fully aligned with either the district or the state standards, that emphasizes fragmented and discrete skill acquisition, and that lacks validation for the types of decisions (e.g., probation and grade retention) made on the basis of its results. Validity with respect to measurement of the goals (e.g., standards) is a critical aspect of an

assessment's quality: if the assessment does not measure what it purports to measure, it could actually draw attention away from the goals of the system rather than toward them. This potential problem is compounded in a situation like that in Chicago, where the use of a single measure and the attachment of consequences to that measure (see below) intensifies attention to the measure rather than to the larger goal of increased student learning. In this regard, it is important to note that the Iowa Test of Basic Skills has not been validated for the purpose of either school or student accountability.

A second limitation of Chicago's and most other school accountability systems concerns the periodicity and specificity (grain size) of the information provided by the outcome measure.<sup>29</sup> On the one hand, a test given once a year that reports a general indication of the content and skills that students have and have not mastered can be extremely valuable for identifying schools and subject areas that may need additional attention, resources, or, possibly, changes in strategies. An important contribution of school accountability systems in places like Chicago and Maryland, for example, is that they have directed the spotlight at failing parts of the system that can then be given additional assistance.<sup>30</sup> However, while such information is useful at these higher levels of aggregation, its potential for directly improving strategies in the classroom is limited. Such assessments are usually administered in the spring to measure student learning during the academic year, but the results are not available in time for the relevant teacher to alter instruction in response to the test. Even if the scores were available earlier in the year, the infrequency and lack of specificity of results is still a problem.

In short, the measure of outcomes through such a test is simply too distant from the complexities of instructional inputs for the teacher to make reasonable attributions of causality. Superstitious learning is common under such circumstances (Levitt & March, 1988). For this reason, some schools supplement the annual testing with more regular and focused assessments aligned with the ITBS. These periodic tests can give teachers more information on outcomes, though they may also underscore concerns about teaching to the test.

A third limitation of the quality of the information is an extension of the second, and concerns the appropriate balance of information about outcomes and information about processes. The implication is that the actor whether that actor is an individual or an organization - must have valid and reliable information on both outcomes and processes. Yet, school accountability systems focus almost exclusively on outcomes, producing little in the way of reliable information on instruction or organizational practices. Some authors have argued that the production of such process-based information at aggregated levels would introduce further measurement problems and unduly constrain practice.<sup>31</sup> That the external accountability system does not generate information on practice might not be a problem if such information were available at the school level. Substantial research suggests that when teachers share information about instruction as well as student learning, they are better able to adapt their practice to the needs and progress of their students.<sup>32</sup> However, the egg-crate structure of U.S. schools impedes such adaptation; information on instructional strategies and processes is held privately by teachers and only rarely shared across the school as the basis for future learning (Lortie, 1975; McLaughlin & Talbert, 2001). Our research in Chicago suggests that bureaucratic school accountability policies are insufficient to establish the patterns of interaction that might foster more effective information sharing in low-capacity schools.

#### -Patterns of Interaction

Indeed, if CPS is any indication, bureaucratic school accountability mechanisms serve to maintain interaction patterns that foster compliance and hierarchy over system learning. Our data reveal a fairly unidirectional (topdown) flow of information throughout the system. For example, rather than being opportunities for collective sharing of information and knowledge, meetings between assistance providers and central office staff and between principals and district liaison personnel were reportedly occasions in which schools and those working with them were simply the recipients of information and mandates rather than sources of valuable information in their own right. When information did flow the other way (from schools and those working with them up into the system), it focused on whether people were carrying out prescribed tasks - that is, whether external partners were providing agreed-upon services, whether schools were implementing specifics of school plans, and whether teachers were understanding and using the tools and techniques disseminated by the external partners (Finnigan, O'Day, & Wakelyn, 2001).

Even those instances in which one might expect more collective problemsolving - such as in the school improvement planning process - more often than not became symbolic exercises in responding to formulaic requirements of the district office rather than thoughtful and inclusive learning experiences for the staff.<sup>33</sup> The planning template was handed down from the Office of Accountability and schools complied, with emphasis on compliance over self-reflection being noticeably stronger in the least improving, lowest capacity schools.<sup>34</sup> What was perhaps most distressing was that this transmission model of information flow also characterized the professional development provided by the external partners, the bulk of which consisted of traditional short workshops rather than intensive inquiry-based explorations of either content or instructional practice (Finnigan et al., 2001).

The end result is that much of the response we saw in schools involved their reacting to directions imposed from above and outside the school rather than reflecting on internal practices. This response is perhaps unsurprising. Hierarchical control and information dissemination are characteristic of large bureaucracies like CPS, and well-established, internalized organizational codes are difficult to change (March, 1991). Moreover, the "get tough" theory of action and the urban politics underlying the school probation policy could be expected to exacerbate these tendencies. By defining the problem as low expectations and a lack of effort on the part of school staff, the forces that come from higher in the system and outside the schools seek to push those inside to work harder. The accompanying incentives only reinforce control and enforcement over system learning.

#### -Maladaptive Incentive Structures

While much of the benefit of current school accountability schemes is supposed to be that they provide incentives to motivate improvement, we found that the incentive structures in Chicago actually exacerbated the problem of motivation in some low-performing, low-capacity schools.

The emphasis on negative incentives (stigma of probation, threat of reconstitution) tied to a single measure (ITBS) appears to have resulted in two tendencies that work against long-term improvement. First, attention in these schools became focused not so much on student learning

per se, but on getting off or staying off probation. This goal essentially places adult desires (to remove the professional stigma and avoid administrative scrutiny) over the needs of students (Wei & Evans, 2001). Second, to achieve this goal, probation schools exhibited an emphasis on strategies to produce immediate increases in test scores, often to the neglect of longer-term success. The combination of these tendencies produced a number of dysfunctional practices.

Most common was the emphasis on test preparation in the form of intensive drill and practice to raise student scores. Some schools even redesigned their curriculum not only to reflect the general skills on the ITBS but to align the proportion of time allotted in the curriculum to a given discrete skill with the proportion of test items measuring that skill. In such cases, the test specifications became the curriculum specifications as well. Another common practice was to triage assistance (mostly test preparation) to students scoring near grade-level cutoffs in the hope that, by raising these students' scores slightly, the school could escape probation. These and similar practices suggest the allocation of resources to achieve adult ends (e.g., getting off probation), rather than to meet the greatest student needs.

Such patterns, which have been noted in prior research on high-stakes testing in education, are not uncommon in organizations in crisis.<sup>35</sup> A focused search for short-term strategies to satisfy a specific target is typical when an organization's performance falls below its aspirations or goals (March, 1994; Simon, 1986). When low performance is combined with negative incentives, including a threat to the organization's position or survival (as is the case of school probation), the potential for maladaptive response increases.

A comprehensive review of research on organizational response to threat (Staw, Sanderlands, & Dutton, 1981) uncovered two dominant and often maladaptive patterns, both of which are relevant to the earlier discussion of organizational learning. First, rather than expanding their use of information to find solutions to the problem, threatened organizations and individuals actually restrict their information processing, relying instead on previously held internal hypotheses and expectations. Exacerbated by stress, this reliance produces a rigidity of action rather than an expansion of strategies and adaptation. A second pattern associated with the presence of a threat is centralization of authority, which in the case of schools serves to further enhance the bureaucratic, control-oriented patterns of interaction mentioned earlier (Staw et al., 1981). Both patterns were observed in low-performing schools in Chicago. We can expect individual and system learning to be constrained under these conditions, as well as innovation and examination of existing practice and assumptions.<sup>36</sup>

A second limitation of the incentive structures is the unbalanced reliance on collective incentives. This a problem raised at the beginning of this article regarding the relationship between individual action and collective accountability. With a focus on schoolwide consequences, the policy offers few incentives for individuals to improve their practice. Individual teacher evaluation is not well aligned with either outcome measures or standards of practice likely to produce those outcomes. In many schools there is little accountability for individual teachers at all, and in others, teachers receive little recognition for improving their practice. One revealing example occurred at one of our case study schools in which a teacher who was working hard to improve practice was repeatedly told by her principal that she could not get an "excellent" rating because "if we had excellent teachers, this school would not be on

probation." In other words, until the school as a whole was removed from probation, this teacher could not expect any reward for her individual efforts, no matter what the actual quality of her work. The effect of this proclamation was a sharp decrease in this teacher's motivation and commitment to the school. While we could attribute this result to the actions of a single principal, what becomes clear is that the policy relies on the ability of the principal (or others in the school community) to motivate the individuals on the school's staff. Where the principal is unable to provide such motivation (often the case in low-performing schools), the effect of the policy on individual teachers is likely to be weak, or even negative. Alternatively, where the principal is an effective motivator or where the connections among individuals are mutually reinforcing, the lack of individual incentives may be mitigated by the strong identification of individuals with the group. This may help to explain why probation schools with higher levels of teacher-to-teacher trust, peer collaboration, and collective responsibility improved more rapidly than others in our Chicago study (Gwynne & Easton, 2001). Perhaps other incentives are at work in these schools to motivate individual behavior.<sup>37</sup>

#### -Weak Resource Allocation and Knowledge Development Strategies

As stated earlier, one of the most promising aspects of outcomes-based school accountability is the use of information to direct attention and resources where they are most needed. In Chicago, this reallocation mainly took the form of external assistance to low-performing schools. Our data indicate, however, how the potential effect of this substantial reallocation of resources was mitigated by the low intensity and lack of focus of most of the support actually provided to schools. For example, external partners<sup>38</sup> on average spent only one or two person-days per week in the schools, and with few exceptions their work provided neither a consistent and coherent focus on literacy instruction (the target of the policy) nor a clear strategy for organizational change (Finnigan et al., 2001). The limitations of the assistance may be attributed in part to problems of implementation, such as the weak quality control in the selection of the external partner candidates. In addition, however, the weak specification of the policy with regard to the content or goals of the assistance gave little guidance to either the schools or the support providers themselves about where to concentrate their energies. Such weak specification is common in school accountability policies in other jurisdictions as well. It derives on the one hand from a desire to respond to internal school context and on the other from the policy's emphasis on student outcomes to the neglect of information - or a theory of action - about instruction. The resulting diffuseness of the assistance, however, does little to highlight or solve problems of attribution discussed earlier. Moreover, the policy neglect of other inputs - such as reallocation of human or other resources - also weakens the potential impact.

#### Summary

What does this discussion of the CPS experience add up to in light of the framework and central problems of school accountability outlined earlier? On the one hand, school accountability policies like those in Chicago have clearly helped focus attention throughout the system on student outcomes and have provided data that can be used for targeting resources and assistance where they are most needed - particularly in low-performing schools. On the other hand, this outcomes-based school accountability approach suffers from a number of inherent weaknesses that make it, as it is currently construed, unlikely to effect the deep changes necessary for long-

term improvement, particularly in low-performing, low-capacity schools. Four weaknesses stand out:

1. The problems of validity, periodicity, and specificity in the outcome measures, coupled with inattention to information on instructional practice, make attribution and thus learning at the school or individual teacher level difficult.
2. Most school accountability systems still operate from a bureaucratic control model and thus fail to create the interaction patterns and normative structures within schools that encourage sustained learning and adaptation. Most low-performing schools lack such patterns and structures.
3. Reliance on negative incentives undermines innovation and risk-taking in threatened schools and diverts attention to organizational survival rather than student learning. Moreover, most current incentive structures fail to foster individual motivation or to reward learning and changes in practice that might lead to sustained improvement.
4. Finally, the reallocation of assistance and resources for increasing the capacity of low-performing schools is generally inadequate and weakly specified. Unfocused assistance based on transmission models of learning does little to build the knowledge base needed for valid interpretation of information produced by the system.

While some of these shortcomings are exacerbated by poor implementation, they derive from fundamental assumptions inherent in the design of current school accountability systems. Current approaches have not solved any of the three problems outlined at the beginning of this article: the relationship between collective accountability and individual action; the tension between external and internal sources of control; and the production, spread, and use of information that can help solve problems of attribution caused by the complexity of school organizations. Thus, reliance on bureaucratic forms of accountability, even with better implementation, is unlikely to lead to the kind of improvement desired.

Is there an alternative?

Professional Accountability: Alternative or Addition?

Perhaps the most commonly posed alternative to bureaucratic or administrative accountability in education is that of professional accountability (Adams & Kirst, 1999; Darling-Hammond & Ascher, 1991; O'Reilly, 1996). Professional accountability is rooted in the assumption that teaching is too complex an activity to be governed by bureaucratically defined rules and routines. Rather, like other professions, effective teaching rests on professionals acquiring specialized knowledge and skills and being able to apply such knowledge and skills to the specific contexts in which they work. In mature professions, the requisite knowledge is articulated in professionally determined standards of practice, and professional accountability involves members of those professions assuming responsibility for the definition and enforcement of those standards.

In education, the focus of professional accountability might be described as three-fold. First, it is centered on the process of instruction - that is, on the work of teachers as they interact with students around instructional content (Cohen & Ball, 1999; McLaughlin & Talbert, 2001). Professional accountability thus concerns the performance of adults in the system at least as much as the performance of students.<sup>39</sup> Second, much of the focus of professional accountability concerns ensuring that educators acquire and apply the knowledge and skills needed for effective practice. Knowledge development is front and center. Third, professional accountability involves the norms of professional interchange. These norms include placing the needs of the client (students) at the center of professional work, collaborating with other professionals to address those needs and ensure the maintenance of standards of practice, and committing to the improvement of practice as part and parcel of professional responsibility.

At the system level, mechanisms of professional accountability center on teacher preparation, teacher licensure, and peer review. At the school level, professional accountability rests both on individual educators assuming responsibility for following standards of practice and on their professional interaction with colleagues and clients. Mentoring, collaboration, and collective problem-solving in response to student needs and some form of peer review to ensure quality of practice are all aspects of school-site professional accountability. Advocates for professionally based forms of accountability argue that this approach holds the most promise for the improvement of teaching and, by extension, for the improvement of student learning.

#### The Promise and Limitations of Professional Accountability

The earlier discussion of complexity and organizational adaptation lends some theoretical support for the claims of advocates of professionalism. In particular, professional accountability at the school site would seem to address problems of attribution and motivation more productively than what we have seen from bureaucratically based models.

With respect to the use of information for effective attribution, three aspects of professional accountability seem most pertinent. First, professionalism draws attention both to instructional practice (agents' strategies) and to teachers' collective responsibility for student learning (outcomes). Second, norms of collaboration around instruction enhance patterns of interaction at the school site that allow for the generation and spread of information about both practice and outcomes and the dissemination of effective strategies based on analysis of that information (i.e., data-driven change). This information is naturally more fine-grained and immediate than that accumulated at higher levels of aggregation, and thus the links between specific strategies and their effects are more easily discerned. This is in part because the articulation of both outcome targets and of standards of practice allows for the testing of hypotheses about those links, particularly where experience or results run contrary to expectations (Axelrod & Cohen, 2000; Sitkin, 1992). Finally, more successful attribution is also likely because of the emphasis on professional knowledge and skills, which lay the groundwork for meaningful interpretations of the information available and more meaningful sharing of information with others. Standards of practice, constructed generally across the profession and more particularly within the professional community of the school, provide the cognitive maps for this process of meaning creation (Huber, 1991).

In addition to addressing problems of attribution, professional accountability expands the incentives for improvement, with particular emphasis on the intrinsic motivators that bring teachers into teaching in the first place, a commitment to students (clients) and an identity as an educator (O'Day, 1996). Scott (1998) has delineated three types of incentives in organizations: material incentives such as monetary rewards and job promotion or loss; community/solidarity incentives based on an identity as a member of a community or profession and a desire to maintain or gain position in that community; and purposive incentives, for example, satisfaction from achieving a valued goal such as a learning objective for students. Our work in schools on probation in Chicago suggests that school accountability policies often fail to tap into solidarity and intrinsic purposive incentives, focusing instead on the threat of material sanctions such as reassignment or job loss or on rewards (Finnigan & Gross, 2001). In contrast, research on professional communities of practice note the motivational aspects of membership in those communities and of normative structures that focus on student learning (goal attainment) and professional identity (Darling-Hammond, 1996; McLaughlin & Talbert, 2001).

Beyond the theoretical appeal of professional accountability, a growing body of empirical evidence points to aspects of professionalism as important components of school improvement. Lee and Smith (1996), for example, find a significant positive relationship between student achievement gains and teachers' collective responsibility for students' academic success in high school.<sup>40</sup> Meanwhile, various researchers have pointed to the positive impact on instruction and student achievement of teacher interaction and collaboration in school-based professional communities (Little, 1990; McLaughlin & Talbert, 1993, 2001; Newmann & Wehlage, 1995). Community School District 2 in New York City provides proof of the deep impact professional culture and professional development can have as a strategy for improvement and system management (Elmore, 1997). Finally, as mentioned previously, there is some recent evidence that professionally based aspects of internal school accountability and capacity are essential for a school's ability to respond effectively to outcomes-based accountability (DeBray et al., 2001; Elmore, 2001) and can help explain the differential gains among schools that are targets of accountability policies (Gwynne & Easton, 2001).

#### Limitations of Professional Accountability

Despite its promise, however, reliance on professional accountability alone cannot ensure that all students' needs are addressed. The most obvious limitation of such a strategy is the overall weakness of professionalism and professional accountability throughout U.S. education. Most U.S. schools are atomized structures in which responsibility rests with the individual educator rather than with the collective body or the profession as a whole (Abelmann & Elmore, 1999; Lortie, 1975). Mentoring and collaboration are simply too rare to ensure the information sharing necessary for ongoing organizational learning. In addition, teachers' knowledge and skills are not what they need to be, especially in light of more challenging goals for all students, and professional standards of practice are only beginning to be defined and enforced.<sup>41</sup> Given this situation, U.S. public school systems may well need external incentives and administrative assistance to stimulate the development of professional accountability and attention to learning objectives for all students.

A second limitation of professionalism concerns the problem of equity. Failure of professionalism is most notable in schools serving disenfranchised groups, especially schools in inner cities with large proportions of low-- income students and students of color. Many of these schools have been allowed to languish for years, during which time the profession has not stepped forth to fight for the needs of these students. External administrative accountability is needed to address this failure of professionalism and help ensure equal opportunity.

Tied to the equity limitations of professionalism is the fact that outcomes-- based school accountability is able to address some systemic purposes and needs more readily than professional accountability alone. For example, at more aggregated levels of the system, performance reporting and other outcomes-based accountability tools provide a useful mechanism for managing the resources necessary for instruction and school improvement. Low performance can be a marker indicating areas of greatest need. In addition, monitoring and reporting of outcomes are important avenues for informing the public about the status of the system and the degree to which it is addressing the needs of and providing opportunities to all students. This is particularly true in the United States, where neither the public nor its representatives are ready to fully trust professional educators (Cohen & Spillane, 1993).

#### A Better Way: Combining the Bureaucratic and the Professional

A combination of professional and administrative/bureaucratic school accountability may be most useful to create an environment that would foster long-term school improvement. Such combinations are common in other professions. In medicine, for example, physicians establish and enforce standards of medical practice but hospitals, insurance companies, and sometimes governments pay attention to outcome (as well as process<sup>42</sup>) data to identify and respond to breakdowns and problems in the distribution and management of medical resources.

Similarly, some educational jurisdictions are experimenting with combinations of outcomes-based administrative accountability to identify areas of low performance and professionally based interventions and accountability to foster school adaptation to address these problems. One such example is the recently formed CEO District in Baltimore, Maryland. Maryland established outcome targets based on the Maryland Student Performance Assessment Program, and identifies schools that are the farthest from and declining in relation to those targets. Such schools are deemed "reconstitution eligible." The preponderance of reconstitution eligible schools in Baltimore City was one factor contributing to a change in governance for the district, to a district-state partnership, and to additional funding being directed to the district (reallocation of resources based on outcomes).<sup>43</sup> Once in place, the new leadership of the district collected a group of the lowest performing schools into a special configuration, the CEO District, and turned to the local education fund (LEF) to assist these schools in becoming effective learning organizations.<sup>44</sup>

The resulting intervention incorporates many of the aspects of teacher professionalism discussed above, as well as the characteristics of adaptive organizations. These include a focus on literacy for teacher and student work and for generating and sharing information about that work; ongoing assessments through Running Records and regular evaluation of student work to provide frequent, fine-grained information about learning outcomes; team structures like regular grade-

level meetings to foster teacher interaction, sharing of strategies, and collective responsibility; and ongoing professional development designed to build a "culture of shared learning" (Fund for Educational Excellence, 2001). Professional development, which is central to the intervention in the CEO District, is site based, focuses on literacy standards, and incorporates teachers' professional interchange through collective study, interclass visitations, and common planning. A client-centered focus, which is also central to notions of professionalism, is manifested throughout these endeavors, but especially in the identification of, targeted assistance to, and monitoring of all students reading below grade level. Finally, the CEO District intervention fosters a full range of incentives, from material rewards for principals and teachers for taking on the challenge of and showing progress in these schools to solidarity incentives derived from membership in a profession, and, finally, to the intrinsic purposive rewards of success with students.<sup>45</sup>

An interesting characteristic of this and several similar approaches is the apparent division of labor between the generation and use of information at the state level, which is following a more administrative/bureaucratic model, and that at the local level, which has a strong infusion of professionalism.<sup>46</sup> This division of labor suggests a necessary distinction in the balance of accountability approaches depending on the level of aggregation, and thus on the distance from or closeness to the point of instruction and school change. What seems critical here is that the bureaucratic accountability mechanisms from the more aggregated levels of the system not get in the way of the development of professional norms, structures, and standards at the school site.

The Baltimore CEO District and similar models represent only one approach to combining professional and bureaucratic accountability. What is important to note is the synergistic interplay of professional and bureaucratic accountability. On the one hand, outcomes-based targets for schools and performance reporting are critical for identifying problem areas, for allocating of resources to address those problems, and for monitoring progress. But the real action at the school site and across school sites is on developing professional knowledge through focused assistance on instruction, professional norms, and the professional patterns of interaction necessary for establishing the basis for ongoing organizational adaptation. This combination allows for a more thorough and balanced incorporation of all aspects of the accountability framework discussed earlier. It generates and draws attention to information relevant to teaching and learning (i.e., to both adult and student performance); it motivates individuals and units through intrinsic as well as extrinsic incentives to attend and to use that information; it builds the knowledge base for valid interpretations of information; and it allocates resources where they are most needed.

The combination also addresses the three underlying problems and tension in school accountability described at the beginning of this article. Recall that the first of these problems concerned the interplay between collective accountability for the school unit and the requisite change in behavior of the individuals within that unit. The addition of professional accountability at the school site strengthens the links between individual teachers and their schools by fostering interaction around common work, a sense of shared purpose, and identity as members of the school community. These ties increase individual motivation to act in accordance with the community's collectively defined endeavor. Moreover, the deprivatization of practice through

sharing of student work and teacher strategies, as well as interclass visitations, provides a mechanism for developing and enforcing common standards of practice.

This latter point also has implications for the second problem identified earlier, the relationship between external and internal mechanisms of control. Attempts to control individual and group behavior by means of external rules and policies are notorious for their inevitable failure, especially in situations where tasks and environments are complex and ambiguous.<sup>47</sup> Resistance and superficial compliance are the common responses. In the case of education, compliance with externally produced rules can even be counterproductive, as it does not allow for the flexible application of professional knowledge to specific contexts and students. In contrast, the strong professional norms generated by the infusion of professional accountability, especially collective responsibility for student learning, become potential resources and mechanisms for orienting the entire school community toward the higher levels of student performance sought by reformers and the general public.

Finally, with respect to the third problem - the generation, flow, and use of information for organizational learning and adaptation - the combination of professional and outcomes-based school accountability holds considerable promise. Drawing on the discussion of complexity theory and the examples described above, we might posit several principles of information generation and use for accountability purposes. The first of these, which is the subject of considerable scholarship and debate, is that information on performance must be valid and accurate and must reflect the goals of teaching and learning.<sup>48</sup> This principle applies to all forms of accountability, regardless of target or mechanism.

With respect to school accountability and improvement, four additional principles derive from our discussion here. First, for individuals and systems to evaluate performance, make appropriate attributions, and adapt their strategies, information must be available both on student performance (achievement) and on adult performance (instructional and other relevant strategies). Second, the grain size and periodicity of the information feedback should match the level and purposes of its use. For the improvement of instructional practice at the school and classroom level, fine-grained and frequent information, including instructionally integrated diagnostic assessments of student learning and feedback on instructional practice tied to that learning, provides the basis for professional reflection. Meanwhile, at higher levels of the system, more aggregate and less frequent information feedback provides a sufficient basis for allocating resources and evaluating and refining policies. Third, because information in complex systems derives from interaction, accountability systems should foster connections within and across units to allow access to and reflection on information relevant to teaching and learning. And, finally, accountability systems must pay particular attention to developing the knowledge base necessary for valid interpretation of the information so generated.

A thoughtful combination of outcomes-based school accountability and professional accountability provides the means of addressing all these information needs and thus for fostering the data driven improvement sought by many system reformers and policymakers. Whether such a thoughtful combination is likely to come about, however, particularly in light of the highly prescriptive and stringent testing and accountability provisions of the No Child Left Behind Act, is another matter.

JENNIFER A. O'DAY

University of Wisconsin-Madison

JENNIFER A. O'DAY is an Assistant Professor at the University of Wisconsin-Madison. She is interested professionally in organizational theory and change, educational equity, and standards-based reform. She is coauthor, with B. Gross, of *One System or Two? Title I Accountability in the Context of State and Local Systems* (1999) and, with S. F. Fuhrman, of *Rewards and Education Reform* (1996).

Footnotes:

1. Work for this paper was supported in part by Office of Educational Research and Improvement (Grant No. R308A60003) to the Consortium for Policy Research in Education and by two grants from the Spencer Foundation (one to the Wisconsin Center for Educational Research for the study of school probation in Chicago elementary schools and the other to Marshall S. Smith to explore implementation issues in standardsbased reform). All findings, opinions, and conclusions expressed in this paper are those of the author and do not necessarily reflect the views of any of the funders.
2. The Consortium for Policy Research in Education unites five leading research institutions - the University of Pennsylvania, Harvard University, Stanford University, the University of Michigan, and the University of Wisconsin-Madison - in collaborative efforts to improve educational systems and student learning through research on educational reform, policy, and finance.
3. The No Child Left Behind Act (NCLB), the most recent reauthorization of the Elementary and Secondary Education Act (ESEA), requires annual testing of all students in grades three through eight, a twelve-year timeline to achieve universal proficiency in reading and mathematics, identification of low-performing schools based on aggregated and disaggregated test scores, and administration of severe consequences (including public school choice) for low-performing schools that fail to improve.
4. Chicago's school accountability policy is explained in greater detail in the section on bureaucratic accountability and in Note 22. In brief, the Chicago system identifies lowperforming schools for "probation" based on student test results and metes out both assistance and sanctions to spur improvement in those schools. This article draws on data from a three-year study of school probation policies and practices in Chicago elementary schools.
5. While this article focuses on the explicit organizational improvement goals of school accountability, these policies have symbolic and political purposes as well. For a fuller discussion of some of the politics underlying Chicago's school probation policies, see Bennett (2001).

6. For a range of implementation discussions over the past three decades, see, for example, Berman and McLaughlin (1974), Goertz, Floden, and O'Day (1995), McLaughlin (1987), Spillane (2000), and Weatherly and Lipsky (1977).
7. See, for example, DeBray, Parson, and Woodworth (2001), Elmore (2001), McLaughlin and Talbert (1993), or Newmann and Wehlage (1995) for a discussion of the power of internal norms.
8. For popular accounts of complexity theory in the natural and social sciences, see Kauffman (1995), Lewin (1992), and Waldrop (1992).
9. See also McLaughlin and Talbert (1993) and Cohen (1990).
10. Interaction patterns among agents within the system are thus critical in understanding how information from the outside becomes available (or not) and used (or not) by its members for learning.
11. "Superstitious learning occurs when the subjective experience of learning is compelling but the connections between actions and outcomes are misspecified" (Levitt & March, 1988, p. 325).
12. It is important to note that organizational inertia is enhanced as the web of supporting relationships grows, creating codependencies among units of the organization and between it and external systems. For a discussion of how initial advantage and increasing returns serves to lock in certain solutions and strategies, see Arthur (1989) or Marion (1999).
13. The term Christmas tree school is often used to indicate a school that adds on multiple uncoordinated and inconsistent programs, much as a Christmas tree displays a melange of multicolored and decorative ornaments.
14. Coupling, in organizations literature, refers to the connections and interdependence of elements of a system. According to Weick (1976), loosely "coupled events are responsive [to one another], but each event also preserves its own identity and some evidence of its physical or logical separateness ... and their attachment may be circumscribed, infrequent, weak in its mutual effects, unimportant, and/or slow to respond.... Loose coupling also carries connotations of impermanence, dissolvability, and tacitness, all of which are potentially crucial properties of the 'glue' that holds organizations together" (p. 3).
15. Axelrod and Cohen (1999) identify three types of attribution mistakes: crediting or blaming a part when a larger ensemble of forces is responsible for a given outcome; crediting or blaming a particular ensemble of factors when a different ensemble is actually responsible; and failing to appreciate the role of context - that is, believing that what was successful in one context may be equally successful in another. Each type of mistake is common in education at every level of the system.
16. An example of such incentives would be salary increases given to teachers for taking additional course credits whether or not such coursework has any bearing on school

improvement plans and strategies, any effect on teachers' actual instructional knowledge, or any impact on student performance.

17. In addition to school accountability, student accountability - in which consequences for individual students such as graduation or grade promotion are based on standardized measures of academic performance - has become an increasingly prevalent aspect of standards-based reform in the past several years. School accountability remains the lynchpin in most jurisdictions, however, as well as in the No Child Left Behind Act.

18. In the case of federal policies (e.g., Title I), to whom may also refer indirectly to the federal government, but still by way of the state and local education agencies (SEAs and LEAs). Similarly, where authority for a given policy derives from state law (legal accountability) and the courts are the ultimate arbiter, administration remains the responsibility of the superordinate levels of the educational bureaucracy to whom the schools are most directly accountable. From the perspective of the schools, then, these cases are almost indistinguishable from other examples of bureaucratic accountability.

19. Note: In this article, the terms bureaucratic accountability, outcomes-based accountability, and outcomes-based bureaucratic accountability will be used interchangeably with school accountability. This discussion excludes choice systems in which schools are held accountable directly to parents through the market.

20. It may be important here to note that the designation "outcomes-based bureaucratic accountability" differs somewhat from the terms used in prior typologies - in large part because it incorporates two somewhat different approaches to categorizing accountability systems. Adams and Kirst (1999) focus their typology on the differences in who is accountable to whom. According to these authors, "Bureaucratic accountability ensures that the preferences and decisions of organizational leaders govern the work of employees throughout the organization. It is based on the relationship between superiors and subordinates and operates through a system of supervisory control characterized by hierarchical structure, standard operating procedures, and rewards and punishments" (p. 467). Darling-Hammond (1990) agrees, underscoring the importance of standardized rules and procedures for practice, established at the top of the system and followed by teachers and others in schools. "New accountability" models do not entirely fit this description, however, as they loosen the standardization of practice, replacing it with standardization of outcomes (or progress toward the outcomes).

O'Reilly (1996) tries to incorporate this distinction by developing a typology based on the operational assumptions that various approaches to accountability make in order to effect systemic change. She thus identifies four operational "theories" of accountability: performance reporting, changes in governance, market-based approaches, and professional accountability. Performance reporting comes closest to the Chicago Public Schools and similar models in that "performance reporting assumes that the collection and reporting of information on student academic performance can be used to stimulate strategies for improving teaching and learning" (p. 5). However, O'Reilly's category falls short in two respects. First, it overlaps with and is integral to other categories in the typology, confounding its interpretation. Second, it does not distinguish between performance and outcomes. While test scores may be a direct indicator of

student performance, they are only an indirect indicator of adult performance. Campbell and his colleagues underscore the importance of distinguishing between performance (behavior) and the results (outcomes) of that performance if there is to be any hope of making valid attributions between the two (Campbell, McCloy, Oppler, & Sager, 1993). This distinction becomes important in understanding the limitations of current models of school accountability to foster improved instruction. It is for this reason that I have chosen the term outcomes-based bureaucratic accountability.

21. The ITBS is an example of the type of commercially produced, norm-referenced, multiple-choice, timed assessment used in many jurisdictions for accountability, monitoring, and placement purposes.

22. All probation schools are required to develop a corrective action/school improvement plan and literacy plan to organize their improvement efforts. These plans are reviewed at the central office, and their implementation is monitored by the probation manager assigned to the school. Assistance comes in the form of additional resources for external support providers called external partners. If the necessary improvement in test scores is not manifest at the end of a vaguely specified number of years, school personnel may be replaced en masse (reconstitution) or individually through a more extended peer review process (reengineering). Such consequences, however, have rarely been administered.

23. The Comprehensive School Reform Demonstration (CSRD) Program, is a federal effort encouraging the adoption of research-based, whole school reform models, especially in Title I schools identified as being in need of improvement.

24. This does not include the resources allocated for remedial summer or after school classes or for practice and testing materials. See Finnigan, O'Day, and Wakelyn (2001) for a fuller discussion of the assistance program.

25. The most recent reauthorization of ESEA, the No Child Left Behind Act of 2001, tightens and extends these requirements and the consequences for continued low performance (see Note 3).

26. An example is California's Public School Accountability Act of 1999, which metes out monetary awards to all schools in the state (both high and low performers) that reach their annual improvement targets on the state's Academic Performance Index (API).

27. For varying perspectives of the impact of these accountability systems, see Grissmer, Flanagan, Kawata, and Williamson (2000), Haney (2000), Klein, Hamilton, McCaffrey, and Stecher (2000), and Koretz and Barron (1998).

28. This research was conducted by a team of researchers from the University of Wisconsin-Madison and the Consortium on Chicago School Research and was sponsored by the Spencer Foundation and the Office of Educational Research and Improvement.

29. By periodicity I mean the frequency and regularity of information on student learning generated by the system. Specificity refers to the degree of detail of the information. For example, information on student knowledge regarding particular decoding skills

(like facility with beginning or ending consonants) is more specific than would be a single test score covering all aspects of decoding, let alone a single score for reading.

30. The No Child Left Behind Act of 2001 promises to substantially alter the specifics of many existing school accountability systems - for example, the nature of the assessments or the timeframe for improvement - but the general intent (to implement a system that identifies lower performing units for intervention) remains.

31. See, for example, *The Debate on Opportunity to Learn Standards: Supporting Works* (National Governors' Association, 1993) or the "Report of the National Academy of Education Panel on Standards-Based Education" (McLaughlin & Shepard, 1995).

32. See, for example, the literature on the role of professional community (e.g., McLaughlin & Talbert, 1993, 2001; Newmann & Wehlage, 1995) and on information sharing in high-performing schools (Darling-Hammond, 1996; Mohrman & Lawler, 1996).

33. This conclusion derives primarily from our observations and interviews regarding the planning process in our ten case study schools but is augmented by additional analysis of the planning documents and processes in a slightly larger and more diverse group of schools. See Gross, Wei, and O'Day (2002) for further discussion of this pattern.

34. Some exceptions exist, of course, with the more self-reflective school communities generally showing the greater gains in performance (Gross, et al., 2002).

35. See, for example, Firestone and Mayrowetz (2000), Nolen, Haladaya, and Haas (1992), and Smith and Rottenberg (1991).

36. These tendencies might be mitigated if a policy included positive incentives (rewards) for learning and for improvements in instructional practice. While some systems do include rewards, these are tied to improvements in outcomes, not practice or learning. Observers often note the need for interim indicators of organizational practice and a capacity to be included in the accountability structures, but they rarely are.

37. For a discussion of these other incentives, see, for example, Mohrman and Lawler (1996) and Darling-Hammond (1996).

38. External partners are the official external support providers approved and funded through the school probation policy (see Note 22).

39. See Campbell et al. (1993) for the importance of the distinction between performance and outcomes.

40. See also Porter and Brophy (1988).

41. The work of the National Board for Professional Teaching Standards (NBPTS), the Interstate New Teacher Assessment and Support Consortium (INTASC), and the National Council for Accreditation of Teacher Education (NCATE) are promising but still beginning efforts in this regard.

42. An example of process data in the medical field might be the number of patients seen in a day or the average length of hospital stay, indicators with which health maintenance organizations and insurance companies have become noticeably concerned in recent years.

43. The partnership between the Baltimore City Public School System (BCPSS) and the Maryland State Department of Education (MSDE) was established by Maryland State Senate Bill 795 in 1997 and is a unique arrangement in educational governance. The partnership decoupled BCPSS from control of Baltimore's city government and established a new Board of School Commissioners that was jointly appointed by the mayor and the governor. It also set in place certain steps for the new board to take in its reform efforts and provided additional monies to support implementation of those steps. For a description and evaluation of the partnership, see Westat (2001).

44. Local education funds (LEFs) are community-based organizations and coalitions centered on improving public schools and student achievement through community involvement in specific locales. The Baltimore LEF, the Fund for Educational Excellence,

is partnering with BCPSS in the implementation of Achievement First, a whole-school standards-based reform effort focused on improving literacy instruction through professional development of teachers and principals. See Fund for Educational Excellence (2001) and Westat (2001).

45. Teachers, for example, receive an 11 percent supplement to their base salary and seven days of paid professional development at the beginning of the school year. See Kim (2001).

46. Other jurisdictions are also experimenting with combinations of professional and bureaucratic accountability, each of which focuses substantially on intense site-based professional development and efforts to develop professional patterns of interaction and collegiality. These include the Boston Public Schools, Community School District 2 and the Chancellor's District in New York City, and the San Diego City Schools, among others.

47. To paraphrase McLaughlin (1988), you can't mandate what matters.

48. See, for example, Linn (2001, 1997).

#### References:

Abelman, C., & Elmore, R. (1999). When accountability knocks, will anyone answer? (CPRE Research Report Series RR-42). Philadelphia: Consortium for Policy Research in Education.

Adams, J. E., & Kirst, M. (1999). New demands for educational accountability: Striving for results in an era of excellence. In J. Murphey & K. S. Louis (Eds.), *Handbook of research in educational administration* (2nd ed., pp. 463-489). San Francisco: Jossey-Bass.

Arthur, W. B. (1989). The economy and complexity. In D. L. Stein (Ed.), *Lectures in the sciences of complexity* (pp. 713-740). Redwood City, CA: Addison-Wesley.

Axelrod, R., & Cohen, M. D. (1999). *Harnessing complexity: Organizational implications of a scientific frontier*. New York: Free Press.

Bennett, A. (2001, April). The history, politics and theory of action of the Chicago probation policy. Paper presented at the annual meeting of the American Educational Research Association, Seattle.

Berman, P., & McLaughlin, M. (1978). *Federal programs supporting educational change: Vol. VII. Factors affecting implementation and continuation*. Santa Monica, CA: RAND.

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt, W. C. Borman, & Associates (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco: Jossey-Bass.

Chicago Public Schools. (2002). *Iowa test of basic skills: Citywide results over time, 1997-2002* (Report ITOT-CW-white). Chicago: Author. Available at [http://research.cps.k12.il.us/resweb/pdf/itbs-over-read\\_a.p](http://research.cps.k12.il.us/resweb/pdf/itbs-over-read_a.p)

Cohen, D. K. (1990). A revolution in one classroom. *Educational Evaluation and Policy Analysis*, 12, 327-345.

Cohen, D. K., & Ball, D. L. (1999). *Instruction, capacity, and improvement*. Philadelphia: Consortium for Policy Research in Education.

Cohen, D. K., & Spillane, J. P. (1993). Policy and practice: The relationship between governance and instruction. In S. H. Fuhrman (Ed.), *Designing coherent education policy: Improving the system* (pp. 35-95). San Francisco: Jossey-Bass.

Daft, R. L., & Weick, K. E. (1984). Toward a model of organizations as interpretation systems. *Academy of Management Review*, 9, 284-295.

Darling-Hammond, L. (1990). Teacher professionalism: Why and how? In A. Lieberman (Ed.), *Schools as collaborative cultures: Creating the future now* (pp. 25-50). Bristol, PA: Falmer Press.

Darling-Hammond, L. (1996). Restructuring schools for high performance. In S. Fuhrman & J. A. O'Day (Eds.), *Rewards and reform: Creating educational incentives that work* (pp. 144-192). San Francisco: Jossey-Bass.

Darling-Hammond, L., & Ascher, C. (1991). *Creating accountability in big city school systems* (Urban Diversity Series No. 102). New York: ERIC Clearinghouse on Urban Education.

DeBray, E., Parson, G., & Woodworth, K. (2001). Patterns of response in four high schools under state accountability policies in Vermont and New York. In S. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states* (pp. 170-192). Chicago: University of Chicago Press.

Elmore, R. F. (1996). Getting to scale with successful educational practices. In S. Fuhrman & J. A. O'Day (Eds.), *Rewards and reform: Creating educational incentives that work* (pp. 294-329). San Francisco: Jossey-Bass.

Elmore, R. F. (1997). *Investing in teacher learning: Staff development and instructional improvement in Community District 2*. Philadelphia: Consortium for Policy Research in Education and the National Commission on Teaching and America's Future.

Elmore, R. F. (2001, April). *Psychiatrists and light bulbs: Educational accountability and the problem of capacity*. Paper presented at the annual meeting of the American Educational Research Association, Seattle.

Elmore R. F., Abelman, C. H., & Fuhrman, S. H. (1996). *The new accountability in state education reform: From process to performance*. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 65-98). Washington, DC: Brookings Institution.

Evans, L. E., & Wei, H. H. (2001, April). *Focusing the work of teachers and schools: The Chicago public schools probation policy*. Paper prepared for the annual meeting of the American Educational Research Association, Seattle.

Finnigan, K., O'Day, J., & Wakelyn, D. (2001). *Buddy, can you lend us a hand? The provision of external assistance to Chicago elementary schools on probation*. Chicago: Consortium on Chicago School Research.

Finnigan, K. S., & Gross, B. M. (2001, April). *Teacher motivation and the Chicago probation policy*. Paper presented at the annual meeting of the American Educational Research Association, Seattle.

Firestone, W. A., & Mayrowetz, D. (2000). *Rethinking "high stakes": Lessons from the United States and England and Wales*. *Teachers College Record*, 102, 724-749. Fuhrman, S. H. (1999). *The new accountability* (CPRE Policy Brief No. RB 27). Philadelphia: Consortium on Policy Research in Education.

Philadelphia: Consortium on Policy Research in Education.

Goertz, M., Floden, R., & O'Day, J. (1995). *Studies of education reform: Systemic reform, volume 1. Findings and conclusions*. New Brunswick, NJ: Consortium for Policy Research in Education.

Grissmer, D., Flanagan, A., Kawata, J., & Williamson, S. (2000). Improving student achievement: What state NAEP test scores tell us (Report MR-924-EDU). Santa Monica, CA: RAND.

Gross, B., Wei, H., & O'Day, J. A. (2002). Planning for improvement in Chicago schools on probation. Unpublished report, University of Wisconsin-Madison.

Gwynne, J., & Easton, J. Q. (2001, April). Probation, organizational capacity, and student achievement in Chicago elementary schools. Paper presented at the annual meeting of the American Educational Research Association, Seattle.

Haney, W. (2000). The myth of the Texas miracle in education [Electronic version]. Education Policy Analysis Archives, 8(41). Available at <http://epaa.asa.edu/epaa/v8n41/> Huber, G. P. (1991). Organizational learning: The contributing processes and the literatures. *Organizational Science*, 2, 1, 88-115.

Kaufmann, S. (1995). *At home in the universe: The search for the laws of self-organization and complexity* New York: Oxford University Press.

Kelley, C., Odden, A, Milanowski, A., & Heneman, H. (2000). The motivational effects of school-based performance awards (CPRE Policy Brief No. RB-29). Philadelphia: Consortium for Policy Research in Education.

Kim, T. (2001, November 14). News in brief: Some Baltimore teachers set to receive extra pay [electronic version]. *Education Week*. Available at <http://www.edweek.org/ew/newstory.cfm?slug=11briefs.h21>

Klein, S. P., Hamilton, L. S., McCaffrey, D. F., & Stecher, B. M. (2000). *What do test scores in Texas tell us?* Washington, DC: RAND.

Koretz, D., & Barron, S. (1998). The validity of gains on the Kentucky instructional results information system (Report MR-1014-EDU). Santa Monica, CA: RAND.

Lee, V., & Smith, J. (1996). Collective responsibility for learning and its effects on gains in achievement for early secondary school students. *American Journal of Education*, 104 103-147.

Levinthal, D. A. (1991). Organizational adaptation and environmental selection: Interrelated processes of change. *Organizational Science*, 2, 140-145.

Levitt, B., & March, J. G. (1988). Organizational learning. *Annual Review of Sociology*, 14, 319-340.

Lewin, R. (1992). *Complexity: Life at the edge of chaos*. New York: Macmillan.

Linn, R. L. (1997). Evaluating the validity of assessments: The consequences of use. *Educational Measurement: Issues and Practice*, 16, 2, 14-16.

Linn, R. L. (2001). The design and evaluation of educational assessment and accountability systems (CSE Technical Report). Los Angeles: University of California, Los Angeles, Center for Research on Evaluation, Standards, and Student Testing.

Little, J. W. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *Teachers College Record*, 91, 509-536.

Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organizational Science*, 2(1), 71-87.

March, J. G. (1994). *A primer on decision making: How decisions happen*. New York: Free Press.

Marion, R. (1999). *The edge of organization: Chaos and complexity theories of formal social Systems*. Thousand Oaks, CA: Sage.

McLaughlin, M. (1988) Learning from experience: Lessons from policy implementation. *Educational Evaluation and Policy Analysis*, 9, 171-178.

McLaughlin, M. W., & Shepard, L. A. (1995). *Improving education through standards-based reform: A report by the National Academy of Education Panel on Standards-Based Reform*. Stanford, CA: National Academy of Education.

McLaughlin, M. W., & Talbert, J. E. (1993). *Contexts that matter for teaching and learning*. Stanford, CA: Stanford University, School of Education, Center for Research on the Context of Teaching.

McLaughlin, M. W., & Talbert, J. E. (2001). *Professional communities and the work of high school teaching*. Chicago: University of Chicago Press.

Mohrman, S. A., & Lawler, E. E. (1996). Motivation for school reform. In S. H. Fuhrman & J. A. O'Day (Eds.), *Rewards and reform: Creating educational incentives that work* (pp. 115-143). San Francisco: Jossey-Bass.

National Governors' Association. (1993). *The debate on opportunity-to-learn standards: Supporting works*. Washington, DC: Author.

Newmann, F. M., & Wehlage, G. G. (1995) *Successful school restructuring: A report to the public and educators by the Center on Organization and Restructuring of Schools*. Madison, WI: Center on Organization and Restructuring of Schools.

Nolen, S. B., Haladyna, T. M., & Haas, N. (1992). Uses and abuses of achievement test scores. *Educational Measurement: Issues and Practice*, 11 (2), 9-15.

<http://global.factiva.com.ezp2.harvard.edu/ha/default.aspx>

O'Day, J. A. (1996). Incentives and school improvement. In S. Fuhrman & J. A. O'Day (Eds.), *Rewards and reform: Creating educational incentives that work* (pp. 1-16). San Francisco: Jossey-Bass.

O'Day, J., & Smith, M. S. (1993). Systemic school reform and educational opportunity. In S. H. Fuhrman (Ed.), *Designing coherent education policy: Improving the system* (pp. 250-312). San Francisco: Jossey Bass.

O'Reilly, F. E. (1996). *Educational accountability: Current practices and theories in use*. Cambridge, MA: Harvard University, Consortium for Policy Research in Education. Porter, A., & Brophy, J. (1988). *Good teaching: Insights from the work of the Institute*

for Research on Teaching. *Educational Leadership*, 45, 75-84.

Scott, W. R. (1998). *Organizations: Rational, natural, and open systems* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Senge, P. M. (1993). *The fifth discipline: The art and practice of the learning organization*. London: Century Business.

Simon, H. A. (1986). Theories of bounded rationality. In C. B. McGuire & R. Radner (Eds.), *Decision and organization: Volume 2* (pp. 161-176). Minneapolis: University of Minnesota Press.

Siskin, L. S. (in press). The challenge of the high schools. In S. H. Fuhrman & R. F. Elmore (Eds.), *Redesigning accountability systems*. New York: Teachers College Press.

Sitkin, S. B. (1992). Learning through failure: The strategy of small losses. *Research in Organizational Behavior*, 14, 231-266.

Smith, M. L., & Rottenberg, C. (1991). Unintended consequences of external testing in elementary schools. *Educational Measurement: Issues and Practice*, 10(4), 7-11. Spillane, J. (2000). *Cognition and policy implementation: District policy-makers and*

the reform of mathematics education. *Cognition and Instruction*, 18, 141-179. Spillane, J. (2002). Local theories of teacher change: The pedagogy of district policies and programs. *Teachers College Record*, 104, 377-420.

Staw, B. M., Sanderlands, L. E., & Dutton, J. E. (1981). Threat-rigidity effects in organizational behavior: A multilevel analysis. *Administration Science Quarterly*, 26, 501-524. Waldrop, M. M. (1992). *Complexity: The emerging science at the edge of order and chaos*. New York: Simon & Schuster.

<http://global.factiva.com.ezp2.harvard.edu/ha/default.aspx>

Weatherly, R., & Lipsky, M. (1977) Street-level bureaucrats and institutional innovation: Implementing special-education reform. *Harvard Educational Review*, 47, 171-197. Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 21, 1-19.

Westat. (2001). Report on the final evaluation of the City-State Partnership. Rockville, MD: Author.

Document ugher00020021121dya100001