Accountability Policies and Teacher Decision Making: Barriers to the Use of Data to Improve Practice

DEBRA INGRAM
KAREN SEASHORE LOUIS
ROGER G. SCHROEDER
University of Minnesota

One assumption underlying accountability policies is that results from standardized tests and other sources will be used to make decisions about school and classroom practice. We explore this assumption using data from a longitudinal study of nine high schools nominated as leading practitioners of Continuous Improvement (CI) practices. We use the key beliefs underlying continuous improvement—derived from educational applications of Deming’s TQM models—and organizational learning to analyze teachers’ responses to district expectations that they would use data to assess their own, their colleagues’, and their schools’ effectiveness and to make improvements. The findings suggest that most teachers are willing, but they have significant concerns about the kind of information that is available and how it is used to judge their own and colleagues’ performance. Our analysis reveals some cultural assumptions that are inconsistent with accountability policies and with theories of continuous improvement and organizational learning. We also identify barriers to use of testing and other data that help to account for the limited impacts.

Standards and accountability continue to be the dominant policy lever for improving student achievement in America’s schools. Beneath the discussion of who should set standards, the validity of the tests, how best to calculate annual yearly progress, and just who it is that should be held accountable (district, schools, individual teachers, or students) is an unexamined assumption that external data and accountability systems will lead to positive change in the daily interaction between teachers and students.

The assumption is that teachers will try harder and become more effective in meeting goals for student performance when the goals are clear, when information on the degree of success is available, and when there are real incentives to meet the goals.

(Newmann, King, & Rigdon, 1997, p. 43).
For standards and accountability policies to be effective in changing the core technology of education—teaching and learning—schools must use accountability data to make decisions about whether they are meeting standards or not and, if not, then use data to change practices and monitor the effectiveness of those changes. Despite the pivotal role of data use in this and other current school improvement policies, there is little strong empirical research on how these policies affect practice. One recent study of accountability policies in England, Wales, and two American states found, for example, that although the implementation of central assessments influenced changes in what topics were taught, there was little change in teachers' instructional approaches; teachers tended to teach new topics using conventional strategies (Firestone, Fitz, & Broadfoot, 2000). Another study, conducted in Virginia, argues that new teachers like and use standards to inform their teaching practice, but more experienced teachers are resistant and resentful (Winkler, 2002). These findings suggest that policy makers need to know much more about how teachers make decisions about teaching and learning if accountability policies are to have a positive impact on teaching and learning. For example, what types of decisions do teachers make and what types of data are meaningful to them? Or what factors promote or impede teachers' use of data for decision making?

In this article we begin to examine this assumption underlying accountability legislation by exploring the concept of data-based-decision making in U.S. high schools. This article is based on a longitudinal study of high schools nominated as leading practitioners of Continuous Improvement (CI) practices. The purpose of the study is to examine the relationship between school culture and implementation of CI practices. This article focuses on one area of CI practice, data-based decision making, in an attempt to answer the following questions:

1. What can we learn about the culture of data-based decision making in schools?

2. What are the implications of teacher decision making for the implementation of standards and accountability policies?

TEACHERS AS USERS OF DATA: PERSPECTIVES FROM ORGANIZATIONAL LEARNING AND CONTINUOUS IMPROVEMENT/QUALITY MANAGEMENT

Although our study initially focused on high schools that are trying to implement continuous improvement, the significant overlap between the construct of continuous improvement and the construct of organizational learning (OL) makes this paper relevant to the field of research on OL.
Given the increased attention to both as solutions to perceived problems with student achievement, it is important to make this bridge.

The concepts of organizational learning and continuous improvement are both popular in the current educational literature. A late 2002 search of the ERIC system produced 176 items when “schools” and “continuous improvement” were keywords and 313 when “schools” and “organizational learning” (a more recent, but obviously popular, term) were keywords, and the publication date was restricted to the last 4 years. The intersection of the three terms, on the other hand, produced only 14 citations. Similar results were found when we used alternative but related terms, such as “quality management” and “learning communities.”

This lack of intersection is, we argue, surprising, because the two terms, while distinct in their origins, have many features in common, including their importation to educational research from the fields of organizational and management studies. On the one hand, continuous improvement and quality management are viewed by many educators as manipulative administrative tools—in part, perhaps, because they are often embedded in legislative language—while the organizational learning perspective has been quickly embraced by educators and is rarely used in policy settings. Our literature review focuses less on the hidden political connotations of each term and more on their meaning as they are discussed in the educational and organizational literature.

ORGANIZATIONAL LEARNING

The term “organizational learning” (OL) captured the attention of public and private sector administrators with the publication of Senge’s (1990) influential book, The Fifth Discipline, although its emergence can be traced to the 1970s (Argyris & Schön, 1974). In education, OL is seen as a powerful process for accomplishing school improvement objectives and a strategy that is particularly useful for educational administrators who wish to work toward long-term renewal rather than quick-fix changes (Petrides & Guiney, 2002). The “generative concept of organizational learning” (Fullan, 1993, p. 6) provides the fundamental justification in theory: Schools that are learning organizations will be able to invent or adapt better solutions to perennial educational problems. Although not all OL is beneficial (March, 1999), Argyris and Schön (1996) argue that high performance, regardless of how it is defined, hinges on the ability to discover new perspectives, gain new understandings, and create new patterns of behavior on a continual basis throughout an organization. Marks, Seashore Louis, and Printy (2001) use survey data to tie school capacities for OL to teacher’s pedagogy and student achievement.
Organizational learning has provided a research framework for examining school improvement over approximately the last decade. Diversity in definitions, rather than rapid convergence, has become the norm (Cousins, 1998). All writers agree, however, that organizations learn in a way that transcends the aggregated learning of their individual members—that is, OL takes place at a group level, even though individuals contribute to it. Engaged in a common activity in a way that is uniquely theirs, the members of an organization learn as an ensemble possessing a distinctive culture (Cook & Yanow, 1993).

Focusing on intellectual, social, and cultural components of the organization, rather than the simple intersection between the individual and the context (Argyris & Schön, 1974; Brown & Duguid, 2000), we define organizational learning as the social processing of knowledge, or the sharing of individually held knowledge or information in ways that construct a clear, commonly held set of ideas (Louis, 1994). This process may be deliberately cognitive, but it more often develops from the accretion of mutual understanding over time in a stable group (Franke, Carpenter, Levi, & Fennema, 2001; Petrides & Guiney, 2002). Central to this conception is that organizational learning involves turning facts or information into knowledge that is shared and that can be acted upon. Senge (1990) provides a similar perspective, emphasizing such learning organization characteristics as systems thinking, shared mental models, team-based learning, and shared vision building.

Central to the concept of OL in most definitions are (1) learning from past experience, (2) acquiring knowledge, (3) processing on an organizational level, (4) identifying and correcting problems, and (5) organizational change. An organization that learns works efficiently, readily adapts to change, detects and corrects error, and continually improves its effectiveness (Argyris & Schon, 1996; Silens, Mulford, & Zarins, 2002). Mulford (1998) notes the following:

The avant-garde of educational change theory is the idea that schools be treated and developed as learning organizations which do not pursue fixed plans in pursuit of set goals, but structure and develop themselves so that they and their members can continually learn from experience, from each other and from the world around them, so that they can solve problems and improve on a continuous basis. (p. 616)

This article addresses data-based decision making in schools, a practice which is only part of the organizational learning concept. By conducting empirical research on this practice in schools, we can begin to identify the use of this aspect of OL in schools, the challenges in using this practice, and
the implications for schools attempting to grapple with the mountain of data emerging from new accountability systems.

CONTINUOUS IMPROVEMENT

The business concept of "total quality management" has been replaced in many educators language with terms such as "continuous improvement," "quality assurance," or "knowledge work supervision" (Duffy, 1997) to avoid the impression that it is an attempt to deny agency to teachers and students; Stringfield (1995) prefers the term "high reliability schools." Whatever it is called, the elements of continuous improvement and quality management as espoused by Deming (1986), Juran (1988), and others map well onto the current context of school policy and practice. The notion that schools must be responsive to customer satisfaction and data (Rinehart, 1993; Peck & Carr, 1997; Salisbury et al., 1997) is supported by attempts in most states and countries to make individual school's learning results public. The standards and accountability movement facilitates CI in at least one way: CI strongly emphasizes focusing on data—standards and accountability policies presumably define what the customer wants students to learn and how they want it measured/demonstrated. In the United States, national teacher associations have made significant investments in promoting continuous improvement among their members (Hawley & Rollie, 2002). Early evidence from North Carolina suggests that principals responded to the state initiatives with significant behavior change regarding the use of data for school improvement (Ladd & Zelli, 2002).

With organizational learning—as with most school improvement strategies—CI accumulated an array of definitions as it worked its way from theory to practice. In our study, CI in education is defined as a set of practices and philosophies embodied in the following seven categories: (1) continuous improvement; (2) customer input/focus; (3) data-based decision making; (4) studying and evaluating processes; (5) leadership; (6) systems thinking; and (7) training (Detert, Kopel, Mauriel, & Jenni, 2000).

The idea that facts and data should form the backbone of decision making is covered well by the emphasis in the organizational culture literature on ideas about the basis of truth and rationality. What is seen as rational and true enables designated individuals to make decisions (Reynolds, 1986; Saphier & King, 1985). However, the quality management literature insists that a specific form of rationality—cause-and-effect analysis—be endorsed. In education, this often assumes starting with the effects (student test results) and working backwards to experiment with possible causes in curriculum and instruction (see Supovitz, 2002). While this is consistent with many perspectives on teacher-as-researcher (where data are not limited to tests or to quantitative information), it conflicts with
some views of postmodernist, constructivist practice in education, which view the facts associated with student testing as products of a biased political system (Peters, 1989; Lipman, 2002). However many postmodernist researchers admit to some accommodation between the ambiguous nature of facts and their use to construct new knowledge (Huberman, 1999).

As is readily apparent from the previous brief descriptions, there is considerable overlap in the constructs of continuous improvement and organizational learning, with some significant differences also appearing. Table 1 shows similarities and differences between the two constructs.

**METHODS**

This article draws on a longitudinal study being conducted of practices and culture in nine high schools that are implementing CI approaches (Detert, Schroeder, & Mauriel, 2000; Detert, Seashore, & Schroeder, 2001). These

| Table 1: A comparison of continuous improvement theory and organizational learning theory |
|---|---|
| Practices and Values in Continuous Improvement Theory | Practices and Values in Organizational Learning Theory |
| Collective vision—clear goals implied | No set goals, but collective vision |
| Study processes (later literature emphasizes both process and results) | Study and learn about every aspect of school functioning |
| Causal analysis and scientific method—the basis for using data for decision-making | Accepts broader perspectives on truth and methods |
| Focus on customers' needs | Not stated |
| Improvements possible without additional resources | Not stated |
| Not stated | Importance of learning from the past—reflection and reflective dialogue |
| Collaboration necessary for an effective school | Collaboration is required to turn “facts” into commonly agreed upon knowledge |
| Not stated | Not stated but implied |
| Teacher-held knowledge as critical | Teacher-held knowledge as critical |
| Only occurs in social contexts | May be planned or unplanned |

Can occur as a consequence of both individual and group work | A conscious element of organizational procedure |
high schools were selected as exemplars of CI practice, although we have found that many of them are not as advanced in implementation of CI as we initially thought. We are assessing the organizational culture of the schools, as one of the reasons CI and data-based decision making is not being readily accepted and established in these schools.

SITE SELECTION

Site selection began by identifying a national sample of high schools practicing CI with a view toward trying to affect teaching and learning. A list compiled by the American Society of Quality, plus other school districts identified in published articles and by state departments of education, resulted in a list of more than 200 school districts. Further screening on our part attempted to determine if these schools were really trying to implement CI as a total philosophy and set of actions as intended by its designers—Deming, Juran, and Shewart, and their followers (criteria for subsequent screening is described in Detert & Mauriel, 1997). We began data collection with 10 high school sites that qualified to be part of the long-term study. We later added a few exemplary schools that escaped our initial scan and stopped following several of the original 10 that have lost their focus on CI or their interest in being studied. Our purpose was to better understand the dynamics of school change, to answer how and why questions, and to generate useful hypotheses for future research; therefore, we were less concerned with the size than with the quality of our sample (Yin, 1994).

The nine high school sites described in this article are located throughout the United States. The schools range in characteristics: large and small, public and private, rural and urban. Enrollment in these schools ranges from approximately 350 students to nearly 3,000 students; district size ranges from approximately 1,500 students to 45,000 students (see Table 2). Demographically, the sites in the study range from approximately 50–99% White. There is also a wide range of factors such as percentage of students eligible for free/reduced lunch (2% to over 50%), average standardized test scores, rankings within their respective states, and graduation rates.

DATA COLLECTION

Data collection in each of these schools started in 1996. Each school was visited between 2 and 4 times for data collection purposes between 1996 and 1998. The visits were conducted by one or two researchers and lasted from 2 to 4 days. Early visits focused on the site history and CI practice implementation; later visits focused on the collection of data regarding
Table 2. Characteristics of the high school sites in the study

<table>
<thead>
<tr>
<th>Site</th>
<th>Number of Students in District</th>
<th>Number of Students in High School</th>
<th>Setting</th>
<th>Geographic Location</th>
<th>Year Continuous Improvement Started</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>17,500</td>
<td>620</td>
<td>Rural area</td>
<td>South</td>
<td>1991</td>
</tr>
<tr>
<td>B</td>
<td>1,150</td>
<td>368</td>
<td>Medium-size city</td>
<td>Midwest</td>
<td>late 1980s</td>
</tr>
<tr>
<td>C</td>
<td>1,500</td>
<td>533</td>
<td>Small, rural city</td>
<td>Midwest</td>
<td>1992</td>
</tr>
<tr>
<td>D</td>
<td>2,100</td>
<td>680</td>
<td>Small city</td>
<td>East</td>
<td>1991</td>
</tr>
<tr>
<td>E</td>
<td>7,200</td>
<td>2,400</td>
<td>Suburb of medium-size city</td>
<td>Mid-Atlantic</td>
<td>late 1980s</td>
</tr>
<tr>
<td>F</td>
<td>45,000</td>
<td>1,084</td>
<td>Large city</td>
<td>Midwest</td>
<td>1992</td>
</tr>
<tr>
<td>G</td>
<td>1,650</td>
<td>450</td>
<td>Small, rural city</td>
<td>Midwest</td>
<td>1991</td>
</tr>
<tr>
<td>H</td>
<td>12,000</td>
<td>2,750</td>
<td>Rural turning suburban</td>
<td>Southwest</td>
<td>1992</td>
</tr>
<tr>
<td>I</td>
<td>*</td>
<td>475</td>
<td>Medium-size city</td>
<td>Midwest</td>
<td>1990</td>
</tr>
</tbody>
</table>

*This private school is not formally part of a K–12 district system.

culture. We used interviews to collect information on the early visits and on the later culture-focused visits we used both interviews and focus groups.

The individual culture interviews consisted of a variety of open-ended interview questions culled from a number of cultural researchers (i.e., Hofstede, Neuijen, Ohayv, & Sanders, 1990; Lortie, 1975; Pettigrew, 1979; Schein, 1992; Trice & Beyer, 1984). These questions sought information on what is expected of individuals, how individuals interact, how individuals spend their time, what they consider most important, and what types of groups or cliques exist within the school. Focus group participants were asked to record on notecards and bring to the meeting any critical incidents they could think of that had occurred in the last 5 years at the school or in their district. The researcher conducting the focus group then probed the group about these incidents, attempting to uncover the key values, beliefs, and assumptions at play during the historical events themselves and the current interpretation of why these events were critical. This critical incident (or social drama) approach has been argued to provide the best window for viewing the formation and display of culture (Deal & Kennedy, 1982; Pettigrew, 1979; Schein, 1992).

During the entire period, data were collected through 186 individual practice interviews, 98 individual culture interviews, and 19 culture focus groups with 101 individuals (see Table 3). At each site, interviewing was
Table 3. Number of respondents by site for interviews and focus groups

<table>
<thead>
<tr>
<th>Site</th>
<th>Practice Interviews</th>
<th>Culture Interviews</th>
<th>Participants in Culture Focus Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>5</td>
<td>30</td>
<td>12</td>
</tr>
<tr>
<td>B</td>
<td>12</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td>C</td>
<td>43</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>D</td>
<td>9</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>E</td>
<td>50</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>F</td>
<td>18</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>G</td>
<td>8</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>H</td>
<td>29</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>I</td>
<td>12</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Total Persons</td>
<td>186</td>
<td>98</td>
<td>101</td>
</tr>
</tbody>
</table>

Giving Data = 385

The variation in number of visits, number of researchers visiting, and time spent on site reflects the wide range in the size of the schools being studied (i.e., from 30 to 160 teachers).

focused on the teachers, but the principal, assistant principal(s), one or more central office personnel (including the superintendent), one or more board or community members, and, at a few sites, parents and students were interviewed. Subjects were selected for interviews on a random basis from teachers who had the same free period. In subsequent visits, a new random selection was taken. Administrators never participated in focus groups with teachers, since we wanted to elicit the honest opinions of teachers about culture and practices. With the permission of the respondents, all interviews and focus groups were tape-recorded and transcribed verbatim.

Although this article is based on the qualitative data, we also draw on a CI practice survey that was distributed to all teachers in each of the participating schools. This paper-and-pencil survey was designed to measure their level of use of continuous improvement practices. The results, including a discussion of the measures, have been published elsewhere (Detert, Kopel et al., 2000). The survey contains statements describing continuous improvement practices in seven criterion areas, and teachers were asked to note their level of agreement or disagreement with each statement based on a 5-point scale. Survey data were not available from site G.

ANALYSIS

The transcripts were coded by several members of the research team using an a priori coding scheme to mark examples of positive CI values, negative CI values, and neutral statements. We also used an inductively generated
coding scheme developed from a thorough reading and discussion of several transcripts. All codes were entered into the QSR NUD*IST qualitative analysis software program, which facilitated reliability checking between the three coders and allowed for easy and efficient querying of the codes in the database.

DATA-BASED DECISION MAKING IN SCHOOLS COMMITTED TO CI

In this section, we describe the results of our analyses. In our analysis, we looked for both supporting data and minority statements that were contrary to the value and practice of data-based decision making. Our focus here is not only to support the practice but also to point out any contrary evidence that we have obtained. This approach might help explain the generality of the data-based decision making practices and the context and situations in which they might hold true.

Given the long commitment to CI among the high schools in this study, and the current policy climate of standards and accountability, we expected to find teachers and administrators with strong values and beliefs about the necessity of using systematic data to make decisions. What we found, however, was that although a sizable proportion (approximately 40%) of the remarks contained a description of using systematic data for decision making, an equivalent proportion of the remarks reflected the use of anecdotal information, experience, or intuition to make decisions. A smaller group of remarks (about 15%) described using a combination of some type of systematic data and some type of non-systematic data such as anecdotes or intuition. Some typical descriptions of how teachers used systematic data to make decisions are shown below. In a few cases, people noted how their understanding of a situation changed once they had collected some data:

Data showed me that my original hypothesis, that ISS (in-school-suspension) wasn’t working, was wrong. In fact, 78% of the kids don’t return to ISS after two visits. This taught me that what we need to concentrate on as a system is the 22% that are here 3 or more times. We need to see if these are “special cases” or if they’re something we can do something about.

I spend most of my time requiring people to show me the statistics. I won’t make a decision at any level without statistics. For example, a board member came to me saying she’d gotten a “bunch of complaints about transportation.” I asked her to define “a bunch” and she said “three.” That changed the nature of the discussion immediately. In the past, we might have rushed to act based on anecdotes like this.
In contrast, other teachers described making decisions based on anecdotal information or on intuition and experience. Some typical remarks are as follows:

I live in the community—I see them in church, in the community, at football games. I ask them if in English they felt prepared. Most say "yes" although some have said "no." I've made changes based on that.

I think that once you've been doing a job for a couple of years I think all teachers have instincts and you can kind of like feel when things go right or wrong. Now I know that that's something that's pretty vague, but really I think good teachers have that. Almost like an intuition . . .

In contrast to the decision-making styles illustrated previously, a much smaller group of teacher and administrators described using a combination of anecdotes, intuition, experience, and systematic data to make decisions. For example, one teacher gave the following description of how a decision was made:

We attempted to use a survey assessment of each of the publisher's books—I had each teacher give me some input. [We used a] combination of experience and data and experts from the other fields.

[We used] input data from students on a rating scale, summarized the top two choices and used [this information] with [our] own "intuition."

It is unclear from these descriptions how much consideration they would give to each source if they were discrepant. In the first description, for example, it is not clear how the textbook decision would be made if teacher experience contradicted the survey results. Therefore, in situations where teachers and administrators described using a combination of data types it is difficult to assess the extent to which data-based decision making occurred. In each situation, there were some systematic data available, but the information in our database was insufficient to determine the extent to which the decision was based on systematic data. Further exploration is needed in this area to understand how teachers and administrators make decisions in situations where they are using multiple forms of data.

Because of these initial analyses, we realized the culture underlying data-based decision making in these schools was more complex than we anticipated. To further understand the culture around decision making in these schools, our next step was to look for themes within the four contexts in which we had asked teachers and administrators to discuss decision making during interviews and focus groups. In the next section, we
summarize their discussion within each perspective in order to look for similarities and differences in their use of systematic data for decision making. The four contexts included in our data collection are as follows: judging teacher effectiveness, judging school effectiveness, and use of data in a recent decision made alone or as part of a team and while studying a process that needed improvement.

DECLUSIONS ABOUT TEACHER EFFECTIVENESS

In response to our question "How do good teachers gauge the effectiveness of their teaching?" teachers and administrators were far less likely to mention measures of student achievement than measures such as student behavior and affect in class, student feedback on courses, student success in college, and even student success after college. Some sample comments about how teachers judge their effectiveness illustrate how practice differs from what policy intends:

How they [teachers] treat the students, behave in class. On how much they do, themselves, in class. How they treat the other, I mean, I expect, when my students leave this building they’ve learned some things about—you know, other things in life. How to treat people, how to be accountable, responsible, I think that’s very important.

I think they can see it on the face of their kids and on the quality of the project that the kids are able to turn out. Not the test scores, but, like, when the kids come in and you’re saying something, you can tell if the light bulb’s on or not. Or, like when a kid all of a sudden goes “ah-ha” and turns on, the excitement of the kids in the program. The honesty and integrity the kids can show you while in the room. Cause, like, I don’t have grades and things, but I can tell kind of how a kid acts as far as his self-esteem and stuff in my room.

I feel I’m effective when I get a bunch of green horns at the beginning of the year and then at the end of the year they’re basically turning in their homework, they’re acting decent in school and we have good rapport going so that they feel free to come in and talk to me and say, “I need help.” That’s how I feel I’m effective.

In the interview and focus group settings, a single respondent could list multiple ways in which they judged teaching effectiveness. As a result, further analysis was needed to determine the extent to which the far greater frequency of nonachievement indicators in the data was due to individual respondents listing multiple nonachievement indicators and to what extent
it was due to a greater number of individual respondents listing an indicator other than achievement. In other words, were nonachievement indicators more prevalent because respondents tended to list multiple example of this type of indicator, or were there more respondents that used nonachievement indicators to judge teacher effectiveness?

To explore this aspect, we next analyzed our data by looking at the combination of indicators described by each respondent. From this perspective, we found that slightly more than half of the respondents mentioned looking only at nonachievement outcomes to judge teacher effectiveness. This suggests that the type of data many teachers use to make decisions about their effectiveness does not match the type of data mandated in external accountability policies. However, because individual respondents often described multiple nonachievement indicators, the magnitude of the difference between teachers who make decisions based on achievement indicators and teachers who use other indicators is not as large as we originally thought.

Another result of looking at the combination of identified indicators is that no respondent listed achievement as the sole outcome they would consider in assessing teacher effectiveness. Each person that mentioned some type of achievement measure also mentioned another type of outcome. For example, when asked how teaching effectiveness should be defined and measured, one teacher responded as follows:

By the performance of the kids, on the curriculum. Everyone’s pushing towards the test results. I guess that’s valid to a degree, but there are other things as well. It’s their honesty and integrity in these kids, Are they treating each other and the staff with respect? That’s part of it as well. Can you socialize in ways that are positive rather than negative, and those things aren’t measured. Sometimes I think that’s where the big push should really be, and it’s not.

In these cases the teachers are also looking at test scores but feel that tests do not tell the whole story. This suggests that accountability policies, which rely on achievement indicators as the primary data source for determining teacher effectiveness, are unlikely to provide sufficient information for teachers to assess their practice and make improvements.

Another theme in how teachers and administrators make decisions about teaching effectiveness is a strong tendency to rely on data that are gathered anecdotally rather than systematically and to rely on intuition and experience rather than data. Standardized achievement tests were rarely mentioned, and the use of systematically collected information is low even when we defined it broadly by including data sources such as teacher-developed tests and course grades. Each of these stretch the definition of
data-based decision making as understood in CI; being locally developed, their reliability and validity is probably unknown. The same situation holds for the roughly half of the sample that consider other measures than achievement.

DECISIONS ABOUT SCHOOL EFFECTIVENESS

To look at data-based decision making from a slightly different perspective, we also asked teachers and administrators “How should the effectiveness of this school be defined and measured?” When compared to the discussion of teacher effectiveness, several differences emerged. First, respondents who mentioned achievement as a basis for determining school effectiveness were also likely to mention standardized test data. In contrast, those who cited achievement as the basis for decisions about teacher effectiveness tended to view locally generated measures, such as classroom tests or student grades, as more appropriate indicators. This suggests that although standardized tests are not useful for decisions about teacher effectiveness, they do hold some value for decisions about school effectiveness. Second, and not surprisingly, no one mentioned using student course evaluations for school effectiveness decisions, as they did for teacher effectiveness.

We also analyzed the data on school effectiveness decisions by looking at the combination of indicators described by each respondent. The results were similar to our findings with the teacher effectiveness question. About half of the respondents did not mention any achievement indicators and sometimes listed multiple, nonachievement examples, as we suspected.

Graduates fill out surveys after one year, maybe 5 years, 10 years, stating where they are in life, what they learned the most in school and what they feel they missed the most in school.

We’ve been trying to work on gauging how we do on other data as well, not just how many go to college but then how many go to work in fields that we trained for. . . . As a school in general, feeling, tone, the pride in the school. I think kids have got to feel good about their school and good about what’s going on here.

The other half of the respondents described a measure of achievement as one source of information, and, again, they frequently explained why achievement wasn’t a sufficient measure.

There was, however, one high school in our sample that exemplifies the use of data to make decisions about school effectiveness. This school did benchmarking of comparable schools and came back with some marvelous data. They also started to gather data on their graduates and produced a
profile of student effectiveness after graduation, along with standard comparisons of test scores with other comparable schools. Nevertheless, this school is an exception and only demonstrated exemplary practice in this one context—judging school effectiveness. They did not stand out on other areas of data collection, nor was there evidence of systematic use of the school effectiveness data.

RECENT DECISIONS AND DATA USE

In contrast to the first two contexts that specifically address decisions teachers would make in response to state standards and accountability policies, we also asked teachers to identify recent decisions made by themselves or as part of a group or committee, and what information they used in making the decisions. In this situation, teachers were split almost evenly among decisions based on the following types of information: data collected systematically, anecdotal information or intuition and experience, and a combination of these types of information. When asked more specifically to describe the data they used in studying a process that needed improvement, teachers and administrators were more likely to mention using systematic data than either anecdotal information and intuition/experience or a combination of systematic and anecdotal data.

This analysis reveals connections between how likely teachers are to use systematic data and the type of decision they are making. One possible explanation for this is that teachers who have been involved in a process study tend to be more advanced CI users and therefore are more likely to reflect the CI value of making decisions based on systematic data. Another possibility is that the recent decisions teachers identified were less likely to be a decision made as part of a group than the process study they recalled. Perhaps even committed CI teachers are less likely to collect systematic data to make individual decisions than they are to collect it for decisions made by a group or a process study. This suggests that teachers are not averse to using systematic data for decisions, but that they are more likely to do so when they are making school-wide decisions than individual decisions in their classrooms. Further exploration is needed on how the level of decision—group versus individual—may influence the likelihood that the decision will be based on systematic data.

QUALITATIVE DATA SUMMARY: STANDARDS, ACHIEVEMENT, AND TEACHER DECISIONS

Our findings point to two major disconnects between current education policy and how teachers judge their effectiveness and the effectiveness of their school.
First, about half of the teachers and administrators in our sample judge teacher effectiveness and school effectiveness by some indicator other than student achievement. Although the other half (approximately) said that they use student achievement data to make decisions about teacher and school effectiveness, they then went on to list the limitations of achievement data and described other indicators they consult. Because of the prevalence of this response, we surmise that being dismissive of externally generated achievement data is a cultural trait that teachers learn and pass on to other teachers as the "right way" to think, act, and feel about the use of data.

In addition, the variety of indicators people use to judge teacher and school effectiveness suggests that local stakeholders have yet to reach agreement about how to measure effectiveness. Coming from different cultural, social, and economic backgrounds, it is not difficult to see why stakeholders would differ in their beliefs about what measures and data are important. For example, one teacher said the following:

Do we want to make them a worker or do I want them to replace a guy who's doing something that's ineffective? In my perspective, when I teach higher level kids, I want them to be the person finding a new way to do it and make society better, not continuing at $5.90 an hour and have no benefits and put plastic tops on plastic milk bottles for the rest of their lives and be disgruntled society workers who can't find anything good in society.

This teacher inferred that employment after high school is only a sufficient indicator of success when the job provides some sense of fulfillment and satisfaction. In this case, the owner of the local milk bottling plant may disagree with the teacher. Unless better agreement can be reached among stakeholders on fundamental goals, there will be little agreement on what constitutes meaningful data. Without debate and discussion among all involved parties, it is difficult for the school to learn effectively and accumulate knowledge about its success over time.

The second divergence between current policy and the reality of schools that emerges from our findings is that even when achievement data are considered as indicators of teacher effectiveness, locally developed achievement measures, such as teacher assessments and course grades, are still viewed as more critical than standardized achievement tests or norm-referenced state tests that are usually part of accountability policies. When describing decisions about teacher and school effectiveness, teachers and administrators often rely on anecdotal information, their experience, or intuition rather than on information they have collected in a systematic manner.
THE CONNECTION BETWEEN SCHOOL'S CI PRACTICES AND DATA-BASED DECISION MAKING

To further understand the influences on teacher decision-making, we combined the qualitative data with data from the survey of continuous improvement practice. Our hypothesis in this analysis was that in schools with higher levels of CI culture the teachers would evidence greater use of systematic data for decision making because data-based decision making is a key aspect of continuous improvement.

Overall, in spite of the districts' multiyear emphasis on quality management, including professional development for staff members, teachers report weak agreement with statements describing CI practice, although there is some variation by site. Based on a 5-point response scale where 1 indicates strongly disagree and 5 indicates strongly agree, the highest mean rating among the seven scales was 3.86, which indicates that teachers are uncertain about whether or not the statements describe their school. This uncertainty appears across the areas of CI practice. For example, the mean rating on the training scale ranged from 2.12 to 3.35, and the range of means on the customer focus scale was 2.45 to 3.21. The seven scales were (1) continuous improvement; (2) customer input/focus; (3) data-based decision making; (4) studying and evaluating processes; (5) leadership; (6) systems thinking; and (7) training (Detert, Kopel et al., 2000).

Analysis of variance tests indicate that the difference in means among the sites is statistically significant for each of the seven practice criteria \(p < .0001\). (See Table 4.) However, post hoc tests using the Scheffe procedure show no consistent pattern across the seven criteria, indicating that implementation of CI is so uneven that no one site has the highest level of implementation in all areas of CI practice (not tabled). The dominant pattern, though it does not hold for all seven criteria, indicates that school H—and sometimes A and C—had practice levels significantly higher than the other schools in the study, while school B and school D had practice levels significantly lower than the other schools in the study. Thus, instead of a clear continuum of CI culture across the schools, the data reveal a cluster of two to three schools that tend to be on the high end and two schools that tend to be at the low end.

Thus, in exploring the potential connection between data use and school CI culture, we turned to the qualitative data to see if we could discern any differences in teacher comments between these broadly distinguished groups: the schools that tend to be on the high end and the schools that tend to be on the low end. In conducting this analysis we looked for both qualitative indications of the depth of data use (more intensive use, use explicitly connected to thinking about classroom practices, teacher effectiveness, school effectiveness, etc.) and also counted the number of times that data-based
Table 4: A comparison of practice survey results by site

<table>
<thead>
<tr>
<th>Site</th>
<th>N</th>
<th>CF Mean (SD)</th>
<th>CI Mean (SD)</th>
<th>DBDM Mean (SD)</th>
<th>LD Mean (SD)</th>
<th>SEP Mean (SD)</th>
<th>ST Mean (SD)</th>
<th>TR Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>6</td>
<td>3.08</td>
<td>3.44</td>
<td>3.15</td>
<td>3.36</td>
<td>2.94</td>
<td>3.26</td>
<td>2.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.07)</td>
<td>(.96)</td>
<td>(1.04)</td>
<td>(1.09)</td>
<td>(1.00)</td>
<td>(1.10)</td>
<td>(1.14)</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>2.47</td>
<td>2.74</td>
<td>2.58</td>
<td>2.46</td>
<td>2.5</td>
<td>2.73</td>
<td>2.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.07)</td>
<td>(1.12)</td>
<td>(1.23)</td>
<td>(1.16)</td>
<td>(1.11)</td>
<td>(1.25)</td>
<td>(1.05)</td>
</tr>
<tr>
<td>C</td>
<td>6</td>
<td>3.15</td>
<td>3.41</td>
<td>3.13</td>
<td>3.41</td>
<td>3.0</td>
<td>3.27</td>
<td>2.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.03)</td>
<td>(.96)</td>
<td>(.95)</td>
<td>(.94)</td>
<td>(.91)</td>
<td>(.91)</td>
<td>(1.07)</td>
</tr>
<tr>
<td>D</td>
<td>4</td>
<td>2.45</td>
<td>2.66</td>
<td>2.54</td>
<td>2.6</td>
<td>2.51</td>
<td>2.68</td>
<td>2.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.06)</td>
<td>(1.14)</td>
<td>(1.14)</td>
<td>(1.24)</td>
<td>(1.17)</td>
<td>(1.29)</td>
<td>(1.19)</td>
</tr>
<tr>
<td>E</td>
<td>1</td>
<td>2.93</td>
<td>2.99</td>
<td>2.97</td>
<td>3.04</td>
<td>2.76</td>
<td>3.18</td>
<td>2.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.15)</td>
<td>(1.06)</td>
<td>(1.10)</td>
<td>(1.12)</td>
<td>(1.05)</td>
<td>(1.17)</td>
<td>(1.23)</td>
</tr>
<tr>
<td>F</td>
<td>3</td>
<td>3</td>
<td>3.23</td>
<td>3.04</td>
<td>3.58</td>
<td>2.89</td>
<td>2.97</td>
<td>2.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.06)</td>
<td>(1.10)</td>
<td>(1.08)</td>
<td>(1.15)</td>
<td>(1.08)</td>
<td>(1.14)</td>
<td>(1.19)</td>
</tr>
<tr>
<td>H</td>
<td>2</td>
<td>3.21</td>
<td>3.61</td>
<td>3.26</td>
<td>3.86</td>
<td>3.21</td>
<td>3.19</td>
<td>3.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.06)</td>
<td>(1.04)</td>
<td>(1.02)</td>
<td>(.95)</td>
<td>(1.00)</td>
<td>(1.15)</td>
<td>(1.20)</td>
</tr>
<tr>
<td>I</td>
<td>1</td>
<td>3.21</td>
<td>3.41</td>
<td>3.14</td>
<td>3.51</td>
<td>3</td>
<td>3.06</td>
<td>2.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.98)</td>
<td>(.93)</td>
<td>(1.06)</td>
<td>(.98)</td>
<td>(.98)</td>
<td>(1/02)</td>
<td>(1.16)</td>
</tr>
<tr>
<td>all</td>
<td>6</td>
<td>3.01</td>
<td>3.27</td>
<td>3.05</td>
<td>3.39</td>
<td>2.93</td>
<td>3.12</td>
<td>2.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.11)</td>
<td>(1.09)</td>
<td>(1.09)</td>
<td>(1.14)</td>
<td>(1.06)</td>
<td>(1.15)</td>
<td>(1.23)</td>
</tr>
<tr>
<td>F</td>
<td>77.7***</td>
<td>103.2***</td>
<td>34.5***</td>
<td>151.1***</td>
<td>64.6***</td>
<td>32.4***</td>
<td>110.5***</td>
<td></td>
</tr>
</tbody>
</table>

Note: All means are based on 5-point indices; *** = significant at the .001 level or better. CF = Customer Focus; CI = Continuous Improvement; DBDM = Data-Based Decision Making; LD = Leadership; SEP = Studying and Evaluating Processes; ST = System Thinking; TR = Training

decision making was discussed by teachers. Neither analytic strategy revealed clear qualitative difference between the two groups of schools in their use of systematic data in making decisions. In each case, there were exceptions to the pattern of the schools in the high-end group also describing most often the use of systematic data to make decisions. One school stood out as having more teacher comments about data use in the qualitative data. This school was not consistently higher; however, on the quantitative measures of CI culture, which indicates that other factors are operating in influencing how likely teachers are to make decisions using systematic data.

ADDITIONAL THEMES

In addition to a strong reliance on anecdotal information and intuition, several other themes emerged as teachers and administrators discussed how
they make decisions. Although these themes were not as prominent as the themes explored above, they shed some light on why the value of data-based decision making may not be as strong as we expected in these CI committed schools.

MISTRUST OF DATA

In talking about decisions, several teachers described situations where they felt data were misused or not used by others. In their experience, data could be used as a tool to force a decision that has already been made rather than as information to shape a decision. After these experiences, it is likely that teachers won't trust data presented by someone else and may even be averse to collecting data themselves. Some illustrative comments are as follows:

You could bring him (an administrator) a stack of paperwork documenting your side of the story and it was “well, yes, but what I'm talking about is intangible.”

Again, the perception that we had was that the conclusion was already a foregone conclusion and that the data were then going to be construed in such a way that it made our decision look like it was—I think that was one of the major problems that we had was that we seemed to kind of haphazardly collect data.

These comments suggest that some teachers see a common practice of hiding or distorting data. If there are judgmental consequences and punitive retributions to the misuse of data, a teacher might find it difficult to treat facts as their friends. Teacher comments indicate a need to increase the capacity of administrators to use data in decision making. In addition, one barrier to data-based decision making in schools may be the struggles of school administrators trying new leadership styles that involve teacher participation in data use as part of a formal commitment to accountability and continuous improvement.³

MEASUREMENT CHALLENGES

In reflecting on how they make decisions, many teachers noted how difficult it is to measure the things they want and need to know. For example, many teachers would like to know how successful their students are in adult life, seeing this as an important indicator of whether a student's educational experience was of high quality, yet they question how anyone could ever capture this kind of information. Others mention the difficulty of
determining cause and effect in a complex environment. To the extent that teachers believe it is difficult to measure the things that are important to them, they are unlikely to try to gather data to make decisions in these areas. Some exemplary quotes are as follows:

In an ideal world I think that—I think we should be able to look five years down the line and see how the unemployment rate is in the class and see if the students have jobs that are kind of matching up to the potential that they might have had. I don’t know how you would ever be able to do such a thing, but that’s the end all.

The one difference I think in education than it is in business because we’re dealing with people and you can’t assign the same qualities to all of them, you can’t say this year I might have very easy students to work with, next year I might have very difficult ones, but nothing remains constant and I think that’s where we differ.

LACK OF TIME

Another theme in the data was that teachers don’t have time to collect and analyze information to make decisions, and often they see it as a trade-off between teaching and collecting data:

Am I doing flow charts? No. It wouldn’t be that I’m not using the concept, please don’t think that I have not given time to thinking about it, but when it comes down to now we’re . . . asked to produce flow charts . . . on top of everything else that we’re asked to do [and] you just can’t do it. I think all of us tend to think of our students first. If it comes between helping a student work on something or doing a flow chart, I’m going to choose the student.

This theme is frequently found in teachers new to evaluation. They lack experience with learning from systematic data, so they see data collection as competing with their real job, which is working with students. These findings also suggest that schools, despite a commitment to CI, have not made adequate adjustments to provide teachers with the additional time needed to gather and interpret data for decision making. Teachers do not, therefore, see more formal kinds of data analysis of the kind suggested by continuous improvement models, as a priority.
TEACHER EFFICACY

A few teachers told us their responsibility is only to teach the curriculum. The curriculum may be imposed by state standards or by the school district or by local agreement. In any case, these teachers believed if they delivered the curriculum; then it's up to the students to do the learning. For example, in response to the question about how teachers judge their effectiveness, one teacher said:

Well, I think some teachers gauge it by how well their students are doing in class via grades; some of them gauge it via just attitudes of students, even work ethics, how a student comes to class, and I think some teacher's don't do that very effectively. I think they basically teach their curriculum and if the kids don't get it then tough.

This statement reflects an important cultural value that is challenged by data-based decision making at the school level. In the recent past, teachers could ignore outcome data if they wanted to. They could isolate themselves in their classrooms and teach their subject, in which case there is very little organizational learning. The notion that teachers should, collectively, take responsibility for student outcomes is both recent and controversial.

DISCUSSION: IMPLICATIONS FOR THE THEORY AND PRACTICE OF CI AND OL

At the beginning of this article we posited that the limited intersection between theories of organizational learning and continuous improvement were, in part, due to their different origins: one in administrative recommendations for change and the other in practice-based theories of change. Our analysis illuminates some of these differences and suggests why the OL theories have spread rapidly throughout the practice community in education, while CI remains largely a buzzword among policy makers and administrator. It also suggests some of the limitations of both as models for school reform.

The contrast between CI and OL stem from differences in emphasis on (1) the degree to which clear goals are assumed; (2) deliberate quantitative causal analysis versus serendipitous and multiple ways of knowing; (3) the customer as the benchmark for measuring improvement (CI only); (4) improvement with an environment of stable resources (CI only); and (5) the degree to which individual learning from data is central versus social processing, reflection, and reflective dialogue. Our study suggests that the
realities of school life, even in districts that have made a significant commitment to CI, are largely inconsistent with the CI assumptions.

Secondary school teachers are articulate about the lack of congruence between the goal assumptions of their state's accountability legislation (which focus exclusively on increasing students' academic achievement) and their own and their community's expectations for a much less well-defined product of "effective adults." Using the latter criteria, they include as critical goals moral development, adaptability, career success, and life satisfactions. Because data on these important goals are unavailable, teacher often turn to intuitive and qualitative measures of their success. Although they acknowledge that the absence of hard data is frustrating, they are unwilling to bend fully to the narrower vision. This emphasis on the need to combine hard and soft data when assessing teacher and school performance is consistent with the OL perspective and may help to explain why it is significantly more popular in the education literature than CI.

The notion that improvement should focus on the customer gives way, in schools, to the reality of multiple stakeholders—legislators and state administrators, district administrators, students, and parents. While multiple interested parties are acknowledged in the continuous improvement literature, there is still an assumption that different stakeholders will have congruent (or overlapping) expectations. In fact, teachers argue that even within a school there are no common assumptions about how to measure teaching effectiveness, while external parties also disagree. There may be a collective vision, as assumed by the OL literature, but it is vague and unfocused, even in schools that have had consistent and long-term exposure to state and district accountability programs.

OL theory pays little attention to resources, and CI contends that improvement is possible without new resources, but there is one resource that prevents teachers from fully committing to data-based decision making: time. The organization of teachers' work life, which is largely unchanged by the districts' efforts to initiate continuous improvement, does not address the overload under which most teachers believe that they operate. Even when they feel comfortable with the data analysis skills that they have been taught (and many are not), they do not believe that they have been given permission to do anything but add work to a load that is already impossible. Until the time issue is addressed directly, opportunities for the kind of reflective interaction over data (quantitative or otherwise) will be limited.

The final difference between CI and OL that may help illuminate why the latter has filtered more quickly into the language of educators stems from assumptions about the nature of data and how data are turned into actionable knowledge. CI theory, which is reflected in recent accountability legislation, assumes that data provide both a trigger and a tool for
improvement of individual and group performance. Our interviews suggest, however, that teachers are most likely to mention systematic data collection and use when they are involved in groups looking at school processes. It does not appear that teachers are data-phobic but rather that they don’t have recent experience in working with data to improve specific classroom practices.

While the evidence suggests why OL theory may be a more comfortable fit with educational practitioners, it does not fully explain why there are so few schools that look like learning organizations. A major issue that arises in these cases is data quality. While we can easily accept alternative methodological paradigms to the analytic techniques proposed in CI, any kind of learning is dependent on having good information. These cases suggest that teachers emphatically believe that, although their intuition as “good teachers” may be correct, they don’t have the data that they need to change their practice. The sites located in states with “mature” accountability systems may have had more confidence in the state tests, but they, like their counterparts in states just implementing state accountability, believed the data to be inadequate for their needs. Given the absence of formal data collection, by any method, around many of the important school goals, teachers tended (apologetically) to fall back on singular anecdotes, most of which derived from personal experience. The lack of a meaningful shared database prohibits the kind of reflection that lies at the core of organizational learning, in addition to the individual learning that appeared as an impossible burden. Interestingly, teachers want the long-term focus implied by both CI and OL. However, they are unable to capture it given the constraints that they encounter.

Based on our findings, we see seven barriers to establishing a school culture supportive of data-based decision making. We believe each barrier has implications for schools trying to follow continuous improvement principles and operate as learning organizations. The barriers represent three types of challenges that administrators need to attend to if they wish to foster the use of standards and accountability data in teacher decision-making processes: cultural challenges, political challenges, and technical challenges.

CULTURAL CHALLENGES

Culture is a strong determinant of how teachers use data to judge their effectiveness and it influences the type of data that teachers think is needed. As deeply held values in the organization that are often submerged, culture exerts a powerful influence on the way decisions are made, the way organizations learn, and the data that teachers find meaningful and useful. The four cultural barriers are as follows:
Barrier 1: Many teachers have developed their own personal metric for judging the effectiveness of their teaching and often this metric differs from the metrics of external parties (e.g., state accountability systems and school boards).

Barrier 2: Many teachers and administrators base their decisions on experience, intuition and anecdotal information (professional judgment) rather than on information that is collected systematically.

Barrier 3: There is little agreement among stakeholders about which student outcomes are most important and what kinds of data are meaningful.

Barrier 4: Some teachers disassociate their own performance and that of students, which leads them to overlook useful data.

For example, both CI and OL require a rational approach to decision making, which uses data to identify problems and make improvements. We found that most teachers do not typically rely on data to examine the effectiveness of teaching. Changing this requires not only changes in behavior but also in deeply held values. Often there is little agreement among stakeholders in the schools, which leads to disagreements about the types of data that are useful in evaluating teaching effectiveness. In addition, teachers may disassociate their own performance from outcome-oriented effectiveness, essentially believing that there is only a modest relationship between their efforts and student achievement. In the face of this situation, changing the approach toward data-driven teaching will require changes in culture along with changing practices. We believe failed attempts to make these changes have focused too much on changing practices and behavior without also changing the underlying values and assumptions held by teachers.

TECHNICAL CHALLENGES

There are a variety of technical factors that affect the use of data. The barriers in this category are as follows:

Barrier 5: Data that teachers want—about "really important outcomes"—are rarely available and are usually hard to measure.

Barrier 6: Schools rarely provide the time needed to collect and analyze data.

The inability to get data that teachers view as most important meshes with the cultural assumptions noted in the previous section. Teachers, even
when they accept their state’s testing and accountability system as necessary, don’t view the test data as sufficient. Another problem is data overload: Schools have tons of data, but often the data are in the wrong form or teachers have to spend an inordinate amount of their time to get it. Finally, the lack of time is exacerbated by the absence of time for team learning in most high schools (Louis, Marks, & Kruse, 1996). Our findings suggest that teachers are more likely to use systematic data when they are involved, as a group, in studying a process that needs improvement.

Administrators could encourage the use of systematic data by rethinking the organization of teachers’ work, helping teachers to set priorities that include time for examining data, and building in a realistic emphasis on group activity. Some of these technical factors could be overcome by the development of appropriate information systems and providing more time for team learning, but this has occurred in only a few school systems to date.

**POLITICAL CHALLENGES**

The final barrier focuses on the well-demonstrated nonrational side of decision making in schools:

*Barriers 7:* Data have often been used politically, leading to mistrust of data and data avoidance.

Our findings illustrate the inherently political nature of the educational process and the difficulties in being completely objective about decisions. One might argue that this should make the use of data even more important—to get at the truth—but data can often be used in more than one way. Therefore, the extent of data-driven decision making and OL will depend on the level of agreement and good will among the various constituents of the educational process.

Organizational learning is dependent on a favorable environment for accepting of new knowledge and ideas. A politicized process that does not involve serious deliberation and discussion can serve to distort facts and to make power a more important ingredient of decision-making than data. Data, of course, still have a purpose. But the purpose is changed in the face of an adversarial political process. In this case data are used to support a decision or course of action rather than to uncover problems and to objectively determine the best course of action.

There is a basic distinction here between the CI literature and political realities in decision making. The CI literature addresses process improvement where data are focused on measurements of the process and alternatives that might be adopted for improvement. In an inherently political process data serve a different role, which is to justify a particular
position. The political nature of decision making in education—whether at the state or local level—is therefore offered as a reason that schools may find it difficult to follow the CI paradigm.

**CONCLUSIONS**

To summarize briefly, we selected schools because they were likely to provide positive examples of data-based decision making, but we found that they conformed neither to the CI nor the OL models outlined in this article. Nevertheless, we believe that the real efforts of teachers and administrators discussed here provide a basis for further reflection about the systematic use of knowledge to improve the educational experiences of children.

In this article we have shown how CI and OL are underlying theories that can be applied to gain a greater understanding of the problems that are encountered by teachers in using data to make decisions concerning teaching and learning. The barriers presented explain why data are not used to the extent that they might be and provide directions for further empirical research. The teachers interviewed for our study suggest overwhelmingly that the concept of data-based decision making and continuous improvement is ideal but, under current conditions, is also unrealistic. Given the genuine commitment of the districts to encouraging continuous improvement work in schools, it is difficult to imagine that these are anything but best-case scenarios. We suggest that much work needs to be done to translate theory into practice and that most of this work must be done around creating the conditions that could facilitate systematic internal improvement processes.

The current policy context provides ample opportunity for further examination of how data and data-based decisions can be used to improve schools. Much of the popular and policy discussions about accountability systems and their impact (or lack thereof) focus on the politics of data use and assume a lack of good will and capacity among involved constituents. Rather than discussions about schools that have failed, testing environments that have been manipulated to produce acceptable results, policy makers whose underlying ambition is to destroy public education, or incompetent administrators and teachers, we need to focus more attention on the goals and processes of CI and OL under conditions of accountability. We make this assertion not because we advocate either current accountability legislation or any particular version of CI or OL but because schools must come to grips with the fact that we know more about teaching and learning than we did several decades ago, when most teachers and administrators completed their training, and that this knowledge is useless until it is processed and used in classrooms where children are expected to learn. We
also need to increase public knowledge and debates about systemic barriers to improvement and knowledge use, and the kinds of data that are needed to increase democratic discussions about school goals and performance.

Earlier versions of this paper were presented at the 1999 annual meeting of the American Educational Research Association in Montreal, Quebec, Canada and the 2000 annual meeting of the International Congress on School Effectiveness and School Improvement, in Singapore, Hong Kong. We wish to thank the Bush Foundation of Minnesota and the National Science Foundation (Transformations to Quality Organizations) for their support of this work.

Notes

1 None of the 14 references were published articles, and only one (Cano, Simmons, & Wood, 1998) included significant empirical research.

2 Given the large variation in the number of teachers at each site, we weighted the data before doing the analysis of variance procedure to ensure that each teacher's response would have equal weight.

3 The issue of how trust in administrator's intentions affects teachers' willingness to engage in continuous improvement has been treated in greater depth elsewhere (Louis, Ingram, & Bies, 2000).

References


DEBRA INGRAM is a research associate at the Center for Applied Research and Educational Improvement in the University of Minnesota's College of Education and Human Development. Her research interests include school change, the arts and learning, and teacher professional development.

KAREN SEASHORE LOUIS is a professor in the Department of Educational Policy and Administration at the University of Minnesota's College of Education and Human Development. Her research interests include organizational theory, schools as workplaces, and leadership. Recent publications include "A Culture Framework for Education: Defining Quality Values for U.S. High Schools" with J. R. Detert and R. G. Schroeder (Journal of School Effectiveness and School Improvement, 12, 2001), and "School Improvement Processes and Practices: Professional Learning for Building Instructional Capacity" with J. Spillane (in J. Murphy, Ed., Challenges of
