

## PERSPECTIVE

### **The Fail-Safe Schools Challenge: Leadership Possibilities From High Reliability Organizations**

G. Thomas Bellamy  
Lindy Crawford  
Laura Huber Marshall  
Gail A. Coulter

*Background.* As public policies increasingly hold schools responsible for preventing school failure, experiences of other organizations that must operate with high reliability may be helpful. This article builds on previous studies of high reliability organizations to inquire how their strategies might inform efforts to improve reliability in loosely coupled professional organizations such as schools. *Purpose.* The article draws on features of high reliability organizations to propose a three-part framework for fail-safe schools that includes (a) efforts to improve normal operations, (b) early detection of learning problems, and (c) rapid recovery from those problems, together with the school structures and cultures needed to support all three. Without assuming that the exact strategies used in very different organizations will work in schools, the article explores how high reliability organizations might serve as a metaphor that poses new questions for research, policy, and practice associated with school improvement and student learning.

**Keywords:** *high reliability organizations; fail safe; reliability; normal operations; student learning*

**T**he No Child Left Behind Act of 2001 (NCLB) codifies a significant addition to the public's expectations of schools. Once described as organizations concerned with sustaining social legitimacy by conforming appearances to public expectations (Meyer & Rowan, 1978), schools have been profoundly

affected by the widespread adoption of standards for student learning and related measures of student achievement. Together, standards and assessments have forced schools to become more instrumental, goal-oriented organizations at the same time that they strive to meet community expectations (Elmore, 2000; March, 2000).

The NCLB adds yet another expectation to this mix of social legitimacy and academic achievement. The stakes for failure have been raised so high—both for schools and for students—that *high reliability* has become an important aspect of school success. Schools are now challenged to prevent practically all failures and to close achievement gaps among student groups—in short, to ensure highly reliable learning for all students. As NCLB policies take effect, any single child's failure to progress at expected levels constitutes a significant failure for the whole school. Whatever the faults of the assessments through which student progress is measured (Shepard, 2003), the expectation for highly reliable results—each and every child making adequate yearly progress and achieving at grade level—is reshaping the context of school work.

Although applying ideas from other organizations to the unique environments of schools always is risky, this article explores what educators might learn from high reliability organizations (HROs) as schools face demands for reliable student learning. To do this, we build on Stringfield and Datnow's (2002) discussion of how schools might use high reliability strategies to frame an inquiry into theoretical, practical, and empirical issues associated with adapting high reliability strategies to schools.

In this article, we first provide a brief review of the literature on HROs, highlighting strategies that appear to foster reliability in these organizations. Next, we examine the potential of these strategies for schools through an exploration of both similarities and differences between schools and HROs. Then, to draw possible lessons for schools, we propose a fail-safe schools framework that uses the HRO literature more as metaphor than as model. We conclude with questions for practice and research that might stimulate new perspectives on the continuing challenge to improve schools and to serve better those students who are most at risk of failing.

---

Authors' Note: The authors wish to thank Michael Murphy, Rodney Muth, Marlina Fiol, Ed O'Connor, David Conley, Rob Horner, and three anonymous reviewers for extremely helpful comments on earlier drafts of the manuscript. Correspondence concerning this article should be addressed to G. Thomas Bellamy, Vice Chancellor for Academic Affairs, University of Washington–Bothell, Box 358522, 18115 Campus Way, Bothell, WA 98011-8246; E-mail: [tbellamy@uwb.edu](mailto:tbellamy@uwb.edu).

## CHARACTERISTICS OF HROS

The literature on HROs describes how organizations operate when accidents or failures are simply too significant to be tolerated, where failures make headlines. HROs operate under high-risk conditions (Roberts, 1990) and “take a variety of extraordinary steps in pursuit of error free performance” (Weick, Sutcliffe, & Obstfeld, 1999, p. 84). Think, for example, of electric power grids, air traffic control systems, prisoner confinement or transportation, commercial aircraft maintenance, nuclear power plants, and toxic chemical manufacturing. The public expects fail-safe performance, and successful organizations adjust their operations to prevent failures.

The study of HROs has evolved through empirical investigation of catastrophic accidents, near misses, and organizations that succeed despite very trying and dangerous circumstances. Launched by Perrow’s (1984) analysis of the nuclear accident at Three Mile Island, the literature evolved through discussions of whether such accidents are inevitable, as Perrow suggested, or might be avoided through strategies used by organizations that operate successfully in high-risk conditions (Bierly & Spender, 1995; Roberts, 1990). Although there are some similarities between this literature and research on organizational effectiveness and quality improvement, HROs “have been treated as exotic outliers in mainstream organizational theory because of their unique potentials for catastrophic consequences and interactively complex technology” (Weick et al., 1999, p. 81). Recently, however, the HRO literature includes more systematic efforts to articulate links to other organizational theories and leadership challenges (Fiol & O’Connor, 2003; Weick & Sutcliffe, 2001).

Of course, all organizations make some mistakes. Errors can arise from individual lapses and from weaknesses in procedures, technologies, and management systems (Reason, 2000). Failures can also arise because the organization must adapt to changing and unforeseen circumstances that make normal procedures ineffective (Weick, 1987). What distinguishes HROs is not the absence of errors but the ability to contain their effects so they do not escalate into significant failures.

### Strategies to Achieve High Reliability

The pursuit of high reliability is a long-standing goal of both industrial practice and organizational theory. For much of the past century, that effort rested on the twin strategies of standardization and control (Taylor, 1911). In the pursuit of reliable performance, work processes were routinized so that complex processes and person-machine interactions could be repeated with

little variation. And to ensure that procedures were consistent, organizations centralized control and developed tightly coupled structures (Hannan & Freeman, 1984).

It is interesting that as organizations strive for reliability by standardizing procedures and centralizing authority, the resulting complexity and rigidity create different risks of failure (Perrow, 1984). “Some systems are sufficiently complex to allow the unexpected interactions of failures in such a way that safety systems are defeated and sufficiently tightly coupled to allow a cascade of increasingly serious failures” (Perrow, 1999, p. 150). Weick, et al. (1999) elaborated on this risk by noting that failures often result when organizations with standardized procedures confront unexpected events that render these procedures ineffective.

A consistent theme in the HRO literature is that standardization and control are insufficient to produce high reliability. Organizations that face demands for high reliability need strategies to respond to the unexpected events that occur in all organizations (Wheatley & Kellner-Rogers, 1995). Successful HROs use strategies such as (a) extensive training and simulation to help employees anticipate as many problems as possible; (b) use of redundant assignments, so that backups are available to help with early identification of problems and to create checks and balances; (c) delegation of an extraordinary degree of responsibility to very low-level employees to support rapid responding as unexpected problems arise; (d) active use of multiple information sources to monitor operations; and (e) a combination of system flexibility with tight job specifications, so that goals can be reached in many different ways (Roberts, 1990). Five “hallmarks of reliability” proposed by Weick and Sutcliffe (2001) offered a complementary perspective on reliability-enhancing strategies:

- *Preoccupation with failure.* HROs “treat any lapse as a symptom that something is wrong with the system, something that could have severe consequences if separate small errors happen to coincide at one awful moment” (p. 10).
- *Reluctance to simplify.* “Knowing that the world they face is complex, unstable, unknowable, and unpredictable, [HROs] position themselves to see as much as possible” (p. 11).
- *Sensitivity to operations.* HROs “are attentive to the front line, where the real work gets done. . . . Anomalies are noticed while they are still tractable and can still be isolated” (p. 13).
- *Commitment to resilience.* “HROs develop capabilities to detect, contain, and bounce back from those inevitable errors that are part of an indeterminate world. The signature of an HRO is not that it is error free, but that errors do not disable it” (p. 14).

- *Deference to expertise.* “HROs push decision-making down—and around. Decisions are made on the front line, and authority migrates to the people with the most expertise, regardless of their rank” (p. 16).

Although these and similar lists of reliability strategies (LaPorte, 1996; Rochlin, 1993) provide a general understanding of how HROs approach their fail-safe challenges, the strategies do not represent a simple formula for reliability. As several authors have noted, many of these strategies can either enhance or detract from reliability, depending on how they are applied. For example, although Roberts and Libuster (1993) emphasized the importance of redundancy in creating needed checks and balances, Rijkma (1997) noted that redundancy could also increase risks by making component failures less visible and by creating unwarranted confidence.

Furthermore, high reliability requires more than a list of discrete strategies. Several mechanisms that integrate the reliability strategies are emphasized in various accounts of HROs. Some focus on organizational structures that enable backup systems to operate (Reason, 2000; Rochlin, LaPorte, & Roberts, 1987). Others emphasize an organizational culture that combines shared-decision premises with latitude for individual improvisation (Weick, 1987). And still others emphasize heightened awareness, or mindfulness, associated with use of reliability strategies (Weick & Sutcliffe, 2001).

### **CAN SCHOOLS BENEFIT FROM THE EXPERIENCE OF HROS?**

At one level, many of the strategies associated with HROs are familiar to educators. For example, the HRO emphasis on continuous training, reliance on frontline staff to solve unexpected problems, and extensive use of data are similar to many of the recommended “best practices” in schools (Stringfield & Datnow, 2002). However, many efforts to import management strategies and policies from business confront unexpected barriers in the complex contexts of schools (Arif & Smiley, 2003; House, 1998). Consequently, a closer look at how schools are similar to and different from HROs seems important before lessons are drawn for schools.

#### **Similarities Between Schools and HROs**

HROs develop because the public demands very reliable performance. Increasingly, public policies are creating similar expectations for schools.

*Similar expectations for high reliability.* The NCLB and the state policies it has stimulated require that all students master content and performance standards and that schools provide the support needed for all children to make adequate yearly progress. Previously, schools often were considered successful when most students did well. Along with advocates for under-achieving students, the NCLB provides a different perspective. Even one child who does not achieve proficiency is increasingly considered a significant failure, as are the achievement gaps that have traditionally separated different student groups (Skrla & Scheurich, 2001). Like HROs, schools make headlines when performance on state tests does not meet expectations. Of course, other aspects of schooling, particularly those related to student safety, are associated with public expectations for high reliability. Although these are important, our primary focus in this article is on meeting the newer and less familiar expectation for reliability in academic achievement. Satisfying these demands for reliability is increasingly necessary for public education to maintain legitimacy and public support. The growing presence of alternatives—charter schools, home schooling, and vouchers—underscores the stakes that are tied to reliable performance.

*Similar conflicts between high reliability expectations and broader organizational missions.* Much as in HROs, the public's expectation for high reliability is far narrower than the mission and goals of schools. Assessments reflect only a part of the learning, acculturation, social development, economic preparation, and caring that are expected of schools (Beck, 1993; Goodlad & McMannon, 1997; Secretary's Commission on Achieving Necessary Skills, 1991). One frequently hears concerns from educators that attaching high stakes to annual assessments can narrow the focus of schools and result in insufficient attention to goals that are not directly assessed (Shepard, 2003).

Schools are similar to HROs in this tension between broad organizational goals and specific public expectations for high reliability. HROs must constantly balance their investments in achieving reliability with other organizational needs and goals (Rochlin, 1993). For example, although the public has little tolerance for error in commercial airline safety, the airlines are also concerned with providing travel experiences that satisfy passengers' interests in convenient schedules, comfort, and ability to work in flight. Meeting high reliability expectations often means diverting attention from other areas that might be more immediately related to organizational missions or finding ways to achieve both reliability and other goals.

### Differences Between Schools and HROs

*Differences in organizational structure.* Much of the HRO literature derives from organizations that have complex and standardized procedures and at least initially, function as centralized and tightly coupled organizations. One way to read the literature on HROs is that the apparently successful strategies represent efforts to minimize the risks of failure that are associated with this type of organizational structure (Perrow, 1984). Although some advocates call for the same kind of standardization and centralization in schools, most practice is still quite different, with reasonably autonomous action by teachers who are buffered from external demands by decoupled organizations (Weick, 1976). In addition, schools are quite different from each other, reflecting locally constructed patterns of relationships that develop in the people-intensive work of teaching and learning (Goldring & Greenfield, 2002). Consequently, caution is needed when the strategies that help prevent failure in centralized organizations with standardized procedures are considered for use in schools.

*Differences in decision options when problems arise.* Differences between schools and HROs come into even sharper focus when the organization's response to problems is considered. The HRO literature generally describes a shift in decision making as problems are detected, moving from the standardized procedures associated with normal operations to more individual autonomy and discretion as emerging problems are encountered. Schools, of course, depend more on the individual autonomy of teachers for their normal operations. Simply providing more individual discretion when problems arise is unlikely to be an effective response. Alternatives to normal procedures are needed when problems arise, but they will have to be different from those used by HROs.

### HROs as Metaphor Instead of Model

Although schools face demands for high reliability that are increasingly similar to those confronting HROs, organizational differences cast doubt on whether the *specific* strategies used by HROs would have the same impact if they were transferred directly to schools. Nevertheless, the experience of HROs may still offer important lessons for schools if they are approached as a metaphor, rather than a model, for how schools might respond to demands for high reliability. Like other metaphors, the high reliability literature provides a lens through which schools can be viewed; useful lenses highlight new areas for exploration. "Each time a new metaphor is used, certain aspects of

organizational phenomena are uncovered that were not evident with other metaphors” (Cameron, 1986, pp. 540-541).

Following this reasoning, we ask what organizational functions are uncovered, or brought to the foreground, in studies of HROs? Identifying and analyzing these functions allows us to investigate how schools might address similar functions more systematically, without any assumption that the exact strategies used by HROs will apply. We turn next to just such an analysis. To emphasize that our reflections use the HRO literature as metaphor, rather than model, we have renamed our approach the “fail-safe schools framework,” rather than using the more traditional “high reliability” label.

### **TOWARD A FRAMEWORK FOR FAIL-SAFE SCHOOLS**

The fail-safe schools framework draws on studies of HROs to identify school functions that appear most important in the effort to achieve highly reliable performance. Just as we do not expect the specific strategies used in HROs to transfer easily to the unique environments of schools, we also do not suggest that any single approach to meeting these functions would be useful in all schools. The purpose of the framework is not to prescribe strategies but more modestly, to bring those functions that appear important for reliable performance to the foreground for systematic attention in research and practice.

The starting point for this inquiry is an analysis of the functions that appear to be served by the strategies used in HROs. Reason’s (2000) analysis of how HROs work is helpful: “Error management has two components: limiting the incidence of dangerous errors and—since this will never be wholly effective—creating systems that are better able to tolerate the occurrence of errors and contain their damaging effects” (p. 3). Limiting the incidence of errors is largely accomplished through ongoing efforts to improve normal operations. Containing the damaging effects of errors involves two functions that are distinguished by Weick et al. (1999)—detecting or discovering small problems before they escalate into failures and managing, or recovering from, those problems.

These analyses suggest that three important functions are associated with achieving high reliability: (a) improving normal operations, (b) detecting potential problems, and (c) recovering from those problems. These three functions, together with environmental supports for their use, outline the fail-safe schools framework. As the following sections suggest, a variety of

different educational practices could address these functions. Although these practices are not unfamiliar in schools, neither are they commonplace.

### **Improving Normal Operations**

As one would expect, practically all accounts of HROs include strategies that are focused primarily on improving the organization's normal operations. The fewer errors or failures that are allowed in the first place, the less the organization must do to adjust when problems occur. Improving normal operations includes such strategies as recruiting, training, and supporting individuals with strong technical expertise (LaPorte, 1996), providing intensive and ongoing training (Babb & Ammons, 1996; Roberts, 1990), and developing simulations of rare or unexpected events (Laporte, 1994). Efforts to improve normal operations often involve centralizing control and building tightly coupled structures to support standard use of effective practices.

These efforts to improve normal operations in HROs are similar to many strategies for school improvement. Given the historical pressures for schools to adopt strategies from businesses, it is not surprising that many efforts to improve reliability in response to current policies have involved efforts to standardize school programs and centralize control (Skrla & Scheurich, 2001). Whatever the long-term impact of these approaches might be, research on high-performing schools suggests that a variety of approaches can be successful (Jesse, Davis, & Pokorny, 2004).

A wealth of alternative strategies for improving schools' normal operations is available. For example, the extensive literature on improving teaching and learning through action research (Lytle & Cochran-Smith, 1994), increased use of research-based practices (Coalition for Evidence-Based Policy, 2002), adoption of innovations (Hall & Hord, 1987), professional development (National Staff Development Council, 1995), benchmarking (Baldrige National Quality Program, 2002), and development of comprehensive school reform models (Northwest Regional Educational Laboratory, 2002) all represent strategies to improve the normal operations of schools. These and related strategies engage schools in scanning the larger environment for effective approaches, putting these approaches to use, and learning from these experiences. At their best, these efforts to improve normal operations contribute to organizational learning that improves the capacity of the school to respond to future challenges (Senge et al., 2000).

### Detecting Threats That Can Lead to Failure

Although important for high reliability, improving normal operations is not sufficient. Unexpected threats, human error, and system weaknesses always exist. Many of the strategies used by HROs help the organization identify problems early enough so that they do not escalate beyond repair; they depend on “systems that solicit skepticism” (Perrow, 1999, p. 152), “preoccupation with failure” (Weick & Sutcliffe, 2001, p. 10), and “riveting attention to detail” (Roberts, Bea, & Bartles, 2001, p. 1). HROs develop a “reporting culture” in which workers are encouraged and rewarded for reporting errors or near misses so that the organization is alerted and can make necessary adjustments. “Without a detailed analysis of mishaps, incidents, near misses, and ‘free lessons,’ we have no way of uncovering recurrent error traps or of knowing where the ‘edge’ is until we fall over it” (Reason, 2000, p. 2). Of course, such a culture depends on organizational conditions that make defensive actions and face saving unnecessary or else the threat to individual workers would naturally inhibit reporting (Bain, 1999).

Detecting threats is difficult. As Weick (1987) noted,

Reliability is invisible in at least two ways. First, people often don't know how many mistakes they could have made but didn't, which means they have at best only a crude idea of what produces reliability and how reliable they are. . . . Reliability is also invisible in the sense that reliable outcomes are constant, which means there is nothing to pay attention to. (p. 118)

A recent *New York Times* article makes the point even more vividly with the headline, “A safer runway? Can't tell. No accidents” (2002).

Early detection is supported by redundancy of assignments, because two or more individuals are more likely to notice difficulties than one. Air traffic control workers, for example, work in groups of two so that each provides a check on the other (Roberts, 1990). Detection is also supported by conceptual slack, circumstances in which “a number of diverging theories pertaining to the organization's technology are maintained simultaneously” (Rijpma, 1997, p. 17).

In schools focused on achieving high reliability in annual measures of student learning, detecting problems early means identifying students who are struggling or falling behind soon enough for the school to respond effectively. It is tempting to argue that schools are positioned well to detect problems because they are awash in data. In fact, however, school structures, schedules, and staff skills all limit the effective use of data in most schools. Although fully developed failures may be well documented in annual

assessments, emerging learning problems often go unnoticed, building at the edges until failures are inevitable.

School-based detection involves processes for identifying learning problems quickly enough to intervene before they escalate into failures. Early identification could mean review of data as often as weekly, monthly, or quarterly, depending on the student, subject area, and skill. Although assessment within the classroom helps individual teachers make needed adjustments, data on potential problems must become public within the school for teachers to respond as a larger community. Many different possibilities for early detection can be imagined, and these are illustrated below with four techniques for early detection, stemming from diverse views of teaching and learning: (a) periodic assessments of student learning, (b) teacher referrals, (c) students' monitoring of their own learning, and (d) collaborative analysis of student work.

*Periodic assessment of student learning.* The importance of formative assessments in improving teaching quality is well known (Black & Wiliam, 1998), and a variety of strategies exist to integrate formative assessments into the work of classrooms and schools. For example, curriculum-based measurement (CBM), once exclusively the domain of special education, is now used schoolwide to monitor progress (Germann & Tindal, 1985) as well as predict students' future performance on large-scale tests of achievement (Marston, 2003). CBM was originally conceptualized by Deno (1986) as a tool for teachers to collect data on students' performance in relation to annual curriculum goals. The reliability and validity of CBMs have been established across all grade levels (Shinn, 1998). And because they are easy to administer and score, CBMs can be completed quarterly for every student in a school. When data from these quarterly measures are made public, teachers can identify students who are not making adequate progress. Research demonstrates the accuracy of CBMs in both identifying very young students who may be at risk of future academic difficulties (Good, Simmons, & Kaméenui, 2001) as well as establishing cut scores needed by students to pass high-stakes assessments (Crawford, Tindal, & Stieber, 2001; Marston, 2003).

Other assessment strategies illustrate a range of practical and theoretical approaches. For example, DeFour (2002) described a strategy of periodic formative assessment of all students in a high school as an essential component of a culture of learning. The Hesperia, California School District relies on a tiered approach to detecting potential learning problems in which all students are evaluated three times a year and those who demonstrate the greatest need get the most intensive help (CalSTAT, 2003). Perhaps more important, staff meet weekly to discuss each student's progress and implement a series

of safety nets that identify potential problems before they develop into full-blown failures.

*Teacher referrals.* Teacher judgments can be reliable predictors of later academic and behavioral performance, so the potential clearly exists for teacher-initiated early detection to become a normal part of school operations. Screening processes for children with behavior disorders have long relied on teacher judgment and have demonstrated the validity of teacher referrals (Walker et al., 1988). Teacher judgments have also been shown to be an accurate measure of students' academic achievement (Gresham & MacMillan, 1997; Hoge & Coladarci, 1989). Providing teachers with students' scores on classroom-based assessments strengthens the accuracy of their judgments (Meisels, DiPrima Bickel, Nicholson, Xue, & Atkins-Burnett, 2001), so the combination of classroom-based data and teacher judgment could prove to be a reliable and inexpensive way to detect learning problems.

*Students' monitoring of their own learning.* Self-monitoring and self-evaluation may be promising methods for detecting learning problems and improving academic performance. When self-monitoring, students recognize and record the occurrence or nonoccurrence of behaviors (King-Sears, 1999). During self-evaluation, students compare their performance to a standard, to past performance, or to the performance of others (Zimmerman, 2002). Students can assess their understanding, efforts and task strategies, and the amount of help they need to complete a task (Paris, 2001), and they are able to evaluate their own work in terms of the cognitive and affective features that affect learning (van Kraayenoord & Paris, 1997). Even first graders appear able to evaluate their feelings and academic performance and describe areas in which they are having difficulty (Huber Marshall & Klebe, 2002; Ialongo, Edelsohn, & Kellam, 2001). In addition to direct academic benefits of self-monitoring (Shimabukuro, Parker, Jenkins, & Edelen-Smith, 1999), procedures for sharing this information within the school could make this a viable early detection strategy.

*Collaborative analysis of student work.* The Annenberg Institute for School Reform's protocols for student work provide an additional illustration of how early warnings of student learning problems might be made public in a school (Blythe, Allen, & Powell, 1999). Coaches lead groups of teachers in discussions of student work in various subjects. The discussion is structured to highlight (a) the relationship of the work to content standards, (b) the grade level appropriateness and challenge of the work, (c) the instruc-

tional context that influenced the quality of the work, (d) the relationship between the demonstration of knowledge and the student's ability to express the knowledge in writing, and (e) misconceptions or knowledge gaps that are revealed in the work (Blythe et al., 1999). Collaborative review of student work appears promising as a strategy to identify learning difficulties because it engages several teachers in the evaluation process, stimulates assignments through which students demonstrate their understanding of a subject area, and makes information about problems public within the review group.

### **Recovery From Emerging Problems**

Once problems are detected, organizations that achieve high reliability appear to have both the motivation and the capacity to respond to minor problems in ways that prevent their escalation into failures and accidents. In general, HROs transition to a qualitatively different approach once problems are identified; they "switch from one mode of decision making to another" (Rijpma, 1997). For example, in the U.S. Navy, standard operating procedures aboard an aircraft carrier are quite rigid and redundant. Yet "anyone on an aircraft carrier can call a halt to flight operations if he sees what looks like a dangerous situation" (Roberts, 1990, p. 108), thereby shifting to a backup strategy that fits the particular problem.

As we noted earlier, most of the backup strategies reported in the HRO literature involve shifting from preventing failure through normal operations that are centralized, standardized, or mechanized to containing unexpected events by empowering decentralized decisions and individual judgment. A more general view that helps link this literature to the quite different circumstances in schools is suggested by Reason's (2000) "Swiss cheese model" of reliability:

High technology systems have many defensive layers: some are engineered (alarms, physical barriers, automatic shutdowns, etc.) others rely on people (surgeons, anesthetists, pilots, control room operators, etc.), and yet others depend on procedures and administrative controls. . . . In an ideal world, each defensive layer would be intact. In reality, however, they are more like slices of Swiss cheese, having many holes—though unlike in the cheese, these holes are continually opening, shutting, and shifting their location. The presence of holes in any one "slice" does not normally cause a bad outcome. Usually, this can happen only when the holes in many layers momentarily line up to permit a trajectory of accident opportunity. (p. 3)

Although much of the HRO literature describes a process of shifting from standard procedures to individual problem solving as the recovery strategy,

the opposite can also be imagined. Organizations might also achieve reliability by relying on individual performance for normal operations and gating to more prescribed procedures or collective decision making when problems appear. Such an approach would appear to have particular relevance in traditionally decoupled professional organizations such as schools, which depend on professional judgment for much of the first line of action. Neither individual problem solving nor backup procedures are infallible, but if the “holes” in these different strategies do not line up, the strategies could combine to contain errors.

In any case, effective recovery requires alternatives, and alternatives normally mean a measure of eclecticism, a tolerance for competing views and interpretations. An organization that is always centralized or always decentralized would appear to be far less effective in containing problems than one that can operate in both modes; similarly, one that contains capabilities associated with many different kinds of programs or interventions is more likely to achieve high reliability than one that relies on intensifying a single strategy.

In schools, one can imagine several ways that recovery systems might be designed to respond to early warnings about potential student failure. For example, when potential problems are identified, a school could shift (a) from individual teacher responsibility to group problem solving, (b) from less to more intensive interventions within the same instructional approach, or (c) from one instructional program or orientation to another.

*Recovery through shared responsibility with teacher teams.* Research demonstrates the effectiveness of teacher teams at various stages of recovery. Various prereferral team models exist, including teacher assistance teams (Chalfant & Pysh, 1989) and intervention assistance teams (Myers & Kline, 2001/2002), that are designed to identify strategies for intervening with students before they are at high risk of failing. Although its effectiveness is related to consistency and quality of implementation, this type of teacher teaming has had positive impacts on student learning (Burns & Symington, 2002). Teacher teams also can be effectively used once students have been identified as needing intensive help by pairing general and special education teachers to support content-area learning (Crawford & Tindal, 2002; Nolet & Tindal, 1994). Equally effective and less-structured forms of teacher collaboration to improve student learning have evolved in conjunction with efforts to strengthen professional communities in schools (Louis, Marks, & Kruse, 1996; Marks & Louis, 1999).

For example, teaching is described as a shared responsibility at Highline Community School in Denver, Colorado, where teachers are encouraged to

openly share data on students who are not meeting academic expectations so that these students can spend part of every day receiving instruction in one of multiple programs (ranging from English language-learning programs to gifted programs). This gating into special programs occurs long before any formal (and expensive) academic evaluations are undertaken (Isenburg, 2003).

*Recovery through more intensive interventions.* Recovery from early warnings of academic or behavioral problems can also be achieved by changing the intensity of treatment with individual tutoring, new instructional groupings, or provision of more instructional time. For example, instead of relying on whole-group instruction to teach a very diverse group of learners, teachers might collaborate within and across grade levels to create skills-based mathematics groups where students receive intensive instruction in the area of their greatest need. In the Success for All reading program (Slavin, 1996), students needing the most help receive it in the smallest instructional groups with the most highly trained teachers.

*Recovery through different instructional approaches.* Backup systems can also be constructed to use different instructional strategies and materials from those used by the teachers with frontline responsibility. For example, students who do not show progress when exposed to frontline instructional approaches may benefit from engaging in computer-based learning. Technology-based instruction, once defined by simple drill and practice software programs, has developed into a viable option for students who do not perform well within traditional settings. Interactive and cognitively challenging software, videodiscs, and Internet-based programs now provide students with complex problem-solving tasks and contextualized learning challenges (cf. Cognition and Technology Group at Vanderbilt, 1997). Backup systems can also be composed of specialists in the schools, as would occur when transition to support by a special education teacher, reading specialist, or therapist is managed effectively.

Sugai and Horner (2002) provided an interesting suggestion about how these backup systems might be organized on a schoolwide basis. Their Positive Behavior Support system offers strategies across a continuum to prevent, detect, and remediate behavior problems. Tier 1 of their model represents primary prevention efforts, designed to impede the development of behavior problems through promotion of schoolwide behavior expectations and discipline (normal operating procedures). Tier 2, or secondary prevention, is targeted at small groups of students who do not respond to schoolwide prevention efforts. Tier 3, tertiary prevention, focuses on individual students with

long-standing complex behavior problems for whom individualized instruction, behavior supports, and team-developed interventions are implemented. Kaméenui (1998) described a similar approach for organizing backup systems in reading instruction.

### **Environmental Conditions That Support Detection and Recovery**

In addition to the technical strategies to discover and respond to early warnings of failure, detection and recovery depend on supportive organizational structures and cultures. To inquire into the nature of this supporting environmental context requires some extrapolation from the literature on HROs because of the differences noted earlier between schools and traditionally identified HROs.

Weick and his colleagues (Weick, 1987; Weick et al., 1999) have provided in-depth discussions of the cultural contexts that support traditional HROs. HROs described in the literature typically rely on standardized procedures for normal operations and develop cultures that sanction individual departure from those procedures when threats are detected. Gating from standardized to individually determined responses in times of crisis has two benefits: It allows very rapid responses, and it validates the status and importance of individuals directly involved in the organization's core processes (LaPorte, 1996).

As noted above, schools seldom begin with such standardized procedures and instead, provide for individual discretion as a significant component of normal operations. Consequently, a culture like that described by Weick (1987) would be unlikely to add reliability to the performance of schools. Simply sanctioning more individual discretion is unlikely to help recover from threats that arise from individually determined normal operations. We suggested earlier that schools might achieve high reliability by gating from individually determined procedures in normal operations to more collective approaches to detect and recover from problems. For backup systems to work in this direction, alternative ways are needed to support speed of response and to ensure that loss of individual control in backup systems does not undermine detection and recovery efforts. To meet these needs, we reason that the support system for detection and recovery in schools involves both structural and cultural elements.

*Structural support for detection and recovery.* Structures needed to support detection and recovery include the policy and procedural frameworks to make information on learning problems public within the school, to sanction

different responsibility assignments or decision processes when problems are detected, and to establish the capacity to respond to these early warnings.

When knowledge about potential learning problems is available to only one teacher, recovery strategies are limited to that individual's repertoire. Certainly, individual teacher adjustments to real-time information about learning difficulties are a critical aspect of fail-safe learning. As Weick et al. (1999) noted, however, high reliability involves responding at several levels of scale in an organization—individual adjustments, dyadic interactions, work teams, and organization-wide adaptations. Gating to alternative strategies beyond the single classroom requires structures and policies that make early warnings public within the school.

To enable collective action once problems are detected, schools need a policy structure that establishes when others, in addition to the individual teacher, have sanctioned responsibilities for developing and implementing instructional programs. In most schools, we expect that effective backup systems would have to be structured to function alongside individual responsibility for normal instruction, not create a threat to individual professional decision making. Consequently, a well-developed mechanism for making early warnings public might serve symbolic as well as communication purposes, legitimizing a shift from individual to shared responsibility.

Finally, structures are needed to ensure that alternatives to normal instructional approaches are available in a timely manner when problems are identified. Although informal interactions among teachers can result in recovery systems when strong collaborative cultures exist (Beck & Murphy, 1996; DeFour, 2002), they are likely to be available inconsistently unless structures are in place to support collective work and alternate interventions. Policies and assignments that give grade-level or subject teams, curriculum specialists, or other groups clear responsibilities once early warnings become public appear needed to ensure that a school can gate from individual teachers' instruction to alternatives when problems occur. Of course, it is possible that existing support services in schools, including special and compensatory education and a variety of related services, might provide the needed backup structures if they are organized to respond quickly whenever learning problems are detected.

*Cultural support for detection and recovery.* The strategies and structures associated with high reliability suggest several cultural norms that although not unfamiliar, also are not present consistently in schools. Drawing on studies of the culture of HROs (Bierly & Spender, 1995; Weick, 1987), we frame the needed cultural support as a set of assumptions about the work of teachers that might facilitate use of high reliability approaches in the unique environ-

ments of schools. In each case, the assumption is itself a paradox, a set of opposite beliefs that represent the differences between normal operations and the alternative approaches needed when problems are identified.

The first assumption underlying a fail-safe school culture is a belief that *both teacher autonomy and collegiality are needed to support student learning*. Learning occurs through the interaction of students, teachers, and materials and performance of tasks associated with these interactions (Cohen & Ball, 1998; Doyle, 1983). The sheer volume of these daily instructional interactions means that schools must rely on individual teachers' judgment in relatively autonomous work. Time limitations prevent teachers from discussing more than a tiny fraction of the instructional interactions with others, so logistics reinforce the norms of privacy that are so strong among teachers (Little, 1990). But this respect for the autonomy and privacy of individual teachers limits the learning needed to improve normal operations (Rosenholtz, 1989; Talbert, 1993), undermines detection by making it unlikely that teachers will make public either their own or others' difficulties with particular students (cf. Bain, 1999), and limits recovery possibilities to the instructional repertoires of individual teachers.

Sustaining this pattern of beliefs about autonomy and collegiality is closely related to how responsibility for student learning is framed. If accountability is too tightly linked to individual teacher performance, then the collaboration needed for ongoing improvement, detection, and recovery would be unlikely to occur. Although teachers are responsible for student learning, effective recovery is possible only when their identification of potential learning problems is viewed as a service to the school, not simply as an admission of failure. Normal operations may well depend on teachers working relatively independently and taking individual responsibility for student learning, whereas detection and recovery require a high degree of collaboration and shared responsibility.

The second set of norms underlying a fail-safe school culture is *confidence in the school's programs as well as skepticism that any approach will work with all children*. A school's curriculum, instructional programs, grouping strategies, and schedules reflect cumulative learning about what works in particular contexts. Confidence in—and even advocacy for—these program features is natural and useful. At the same time, however, skepticism about their reliability may also be necessary. The manner in which certain students respond to certain curricular and instructional programs never will be completely stable; unanticipated responses and surprises will always occur, no matter how well developed the programs. Although normal operations may depend on some confidence in the school's programs, detecting problems early may well depend on a more skeptical stance. Weick and

Sutcliffe (2001), building on the work of Langer (1989), described this as a need for *mindfulness*, for constant vigilance that recognizes that problems can occur at any time. A useful mindset that supports this vigilance is an assumption that every program, curriculum, instructional strategy, and teacher is fallible; some students will always fail if a school relies on either packaged programs or individual teacher skills. That is, rather than assuming that a procedure, even a research-based one, or a teacher, even an excellent one, can produce consistently reliable results, it is better to assume the opposite. In Reason's (2000) terms, each is a slice of Swiss cheese that naturally needs backup to catch students who fall through the holes.

The third cultural assumption involves *collaboratively held visions and program models, combined with diversity of ideas about teaching and learning*. Although shared approaches to curriculum implementation have been identified as characteristic of some effective schools (Sammons, 1999), the high reliability literature suggests that the opposite may also be important. It appears more likely that problems will be detected and addressed effectively when individuals with divergent viewpoints work in collaboration to observe and analyze situations. Rather than advocacy for adherence to a single approach, this suggests a more balanced strategy that honors differences within the context of close collaboration.

A fourth cultural norm supporting fail-safe operations *requires that a school's focus on learning be complemented by a preoccupation with failure*. Traditionally, schools have focused most of their attention on maximizing student learning, developing the curricula, instruction, and climate needed to support learning. In spite of the positive focus of this effort, too many students were virtually invisible as their small struggles at 1st grade became increasingly apparent at 8th grade and created drop-out causalities at 10th grade. We tend to notice what we believe we can influence, and a school's determination to ensure success for all students creates, paradoxically, a preoccupation with identifying problems early precisely because they can be addressed through different school approaches. "People preoccupied with failure comprehend more of the potential complex interactions in a system and create alternative paths for task performance" (Weick et al., 1999, p. 105). A preoccupation with failure may increase the school's capacity to notice and respond to learning difficulties, just as a preoccupation with successful learning can stimulate creativity in normal classroom instruction.

### **Features and Limits of the Fail-Safe Schools Framework**

Rather than assuming that the strategies used by HROs would transfer directly to schools, we instead used knowledge about HROs as a metaphor to

identify organizational functions and features that have been highlighted as important to high reliability. The fail-safe schools framework is organized around these functions and investigates possible approaches that schools might take in each area. Our analysis supports the value of metaphors noted by Cameron (1986). Although some strategies associated with each of the highlighted functions are familiar in schools and are included in some program models, they are neither commonplace nor combined into a coherent organizational strategy in the literature on educational leadership.

As an approach to support high reliability in schools, the fail-safe schools framework identifies a set of functions that need to be addressed systematically and simultaneously. The framework proposes that reliability is enhanced when a school (a) engages in continuous efforts to improve normal operations, (b) creates norms and systems to detect problems early, and (c) uses systematic recovery strategies through backup mechanisms and then establishes cultural and structural supports for these activities.

All parts of the framework seem critical to high reliability, for the absence of any one would likely undermine the effectiveness of the others. Early detection is of little use when no recovery strategies are available; backup structures for recovery help little if social norms in the school do not encourage teachers to ask for help and work collaboratively to meet the needs of a diverse student body. And neither detection nor recovery systems are likely to succeed when normal operations produce so many failure indications that recovery systems are overloaded.

The framework is not an argument about how a school should address these four functions. The examples listed in this article are illustrations; they are intended neither as prescriptions nor as an exhaustive list of possibilities. By focusing on functions, rather than strategies, the framework emphasizes that reliability cannot be reduced to a list of techniques or standardized routines. Indeed, one value of the framework may be its invitation for schools to develop varied and eclectic strategies associated with the four functions that respond to today's wide differences in school contexts and approaches to teaching and learning.

We have investigated the implications of the HRO literature for achieving reliability at the level of the individual school. This reflects the increasing focus on the school as the point of public accountability for learning results. Schools' results are published in the newspapers, schools are placed on probation, and sanctions for not meeting adequate yearly progress are placed on schools. Of course, the same four functions might also apply at several levels of scale: for each child, in each teacher's classroom, among teams of teachers, in the school as a whole, and in the operation of a school district. Great teachers continuously improve their practice and create classroom structures

and cultures that support early detection and recovery. Students play a role by monitoring and evaluating their learning and identifying areas of need. Grade-level teams can help bring teachers' collective expertise to bear on detection and recovery challenges. Having fail-safe strategies in place at all levels would probably contribute to reliability in learning outcomes. In the real world, however, it may be just as important for each organizational level to anticipate surprises from other levels that can threaten as well as support reliable learning, increasing the need for constant vigilance to identify and respond to emerging problems.

### **SOME QUESTIONS FOR RESEARCH, POLICY, AND PRACTICE**

Clearly, the framework for fail-safe schools is our own creation. It rests on an effort to extrapolate from studies of organizations that are quite unlike schools, to identify the apparent functions served by features of these organizations, and then to imagine how those functions might be met in schools. Consequently, we frame implications from the review of HROs as a set of initial questions that might be fruitfully explored in educational research, policy, and practice. The underlying issue is straightforward: Does the framework for fail-safe schools add perspectives or questions that might lead to meaningful contributions to the national conversation about school performance? As is clear from the many examples in earlier parts of this article, each of the elements of high reliability theory is familiar in the educational literature, if not consistently present in practice. Our questions relate not to these separate components but to the utility of the complete framework for fail-safe schools.

*Could the fail-safe schools framework contribute perspectives to the challenges of improving school performance?* Leadership for improving school performance is frequently framed as a challenge to align school support efforts so that they provide direct support for the interaction of teachers, students, and materials through which students learn (Spillane & Louis, 2002). Much of the effort to create this alignment depends on the principal's leadership for collegiality and professional community, so that teachers work together in ways that foster both their own learning and that of their students (DeFour & Eaker, 1998; Smylie & Hart, 1999). Although such collaboration can facilitate teacher and student learning (Marks & Louis, 1999), high levels of collaboration do not automatically focus on student learning and its improvement (Elmore, Peterson, & McCarthey, 1996). Framing the challenges

of collaboration around the elements of the fail-safe schools framework might provide an alternative focus for studying professional community. In the language of this article, much of the investigation of professional community has related to teacher and organizational learning that improves normal operations (Elmore, 2000; Spillane & Louis, 2002). Might the effectiveness of leadership for teachers' professional community be enhanced with simultaneous attention to collaboration in detection and recovery? This possibility is consistent with the observation of Weick et al. (1999) that organizational learning—to increase an organization's ability to avoid problems—and a commitment to resilience—to respond to problems when they occur—may provide alternative ways to improve performance and reliability.

*Could the fail-safe schools framework contribute to the discussion of teacher professionalism?* The policy debate about improving schools often contrasts efforts to improve the professional status and responsibilities of teachers with a more technical view of teaching that involves implementing research-based curriculum packages, standard programs, or comprehensive school models (Devaney & Sykes, 1988; Education Commission of the States, 2000). On one side, proponents argue for curriculum materials that are more scripted and standardized, making teaching largely a matter of skilled implementation of programs developed elsewhere (Pogrow, 1996). Those advocating greater teacher professionalism disagree, noting that the contexts of teaching and the needs of children vary so widely that exceptionally skilled teachers are always needed to adapt whatever materials and procedures are used (Darling-Hammond, 1999).

The fail-safe schools perspective does not resolve these differences, but it does argue for a synthesis. At present, we have no evidence that either great teacher education and professional development or the very best research-based practices can, by themselves, produce the level of reliability expected of schools. With our current state of knowledge and today's reliability expectations, neither can succeed without the other. Consequently, a productive line of inquiry might investigate how a combination of both might be structured in policy and organizational design. The notions of detection, recovery, and environmental support provide a possible entry point for this investigation. To shift from one approach under normal operations to a different approach in response to early warnings of learning problems, a school and its teachers need the capacity to use more than one mode of operating. A useful line of inquiry might address how these simultaneous opposites can be sustained in the culture of a school and the profession.

*Might the fail-safe schools framework help in the current effort to implement research-based practices in schools?* From the perspective described in this article, much of the current effort to implement research-based practices can be seen as policy-driven strategies to improve normal operations of schools (Shavelson & Towne, 2002). And although much attention has been given to fidelity of implementation of research-based practices (Desimone, 2002; Haynes, Emmons, & Woodruff, 1998), the fail-safe framework suggests that the results achieved in such implementation might vary considerably, depending on whether a school had other aspects of high reliability in place. It may be that different results associated with implementation efforts are related to the presence of reliability-producing elements in the school as a whole, rather than to more specific aspects of program implementation.

*Can the fail-safe schools framework help explain the wide discrepancies in identification of children eligible for special education services?* Wide differences in placement rates into special education have been noted for some time (MacMillan & Reschly, 1998; Oswald & Coutinho, 1995), and controversy surrounds the process of identification of children with learning disabilities (O'Shaughnessy, Lane, Gresham, & Beebe-Frankenberger, 2003; Peterson & Shinn, 2002). Seldom is the quality of classroom instructional practices considered in the placement of students in special education (Harry, Klingner, Sturges, & Moore, as cited in Artiles, Harry, Reschly, & Chinn, 2002). It is possible that referrals to special education are lower in schools or classrooms that have the major elements for high reliability in place. A report of one school district that appears to have well-developed procedures for detection and recovery supports this possibility. Using an approach labeled *neverstreaming* (a term originally coined by Slavin & Madden, 1989), the Elk Grove, California School District uses strategies that identify and respond to learning problems early; their results include a 44% reduction in special education enrollment (Ortiz, 2001). When early detection and recovery are a normal part of a school's functioning, it may reduce the need to identify individual deficits that create eligibility for special education. If this possibility is supported in research, it would also encourage further development of alternative models for early identification and intervention with children who experience learning problems before they are formally evaluated for special education services.

## SUMMARY

We have used the literature on HROs as a metaphor to uncover organizational functions that seem important as organizations strive to achieve the highly reliable results that are increasingly expected of schools. We use the fail-safe schools framework to explore the application of these functions in schools and to raise questions for theory, policy, and practice as educators strive to meet new public expectations for more reliable student learning. Taken separately, most of the procedures suggested by the dual challenges of detection and recovery are familiar in the educational literature if not commonplace in schools. Furthermore, both detection and recovery strategies are present in various models of teaching, learning, and school organization. As a result, the framework offers a possible lens for studying exemplary school performance and school improvement strategies that cuts across familiar theoretical perspectives. The fail-safe schools framework offers similarly unconventional suggestions for studying some current policy conflicts. Rather than supporting one side or another in the debates about school improvement and professionalism, the fail-safe framework challenges us to find alternative organizations and policies that support more creative combinations of the opposing views. The fail-safe schools framework invites research and testing to see what strategies are most effective under different school conditions and what results actually obtain when each of the fail-safe functions is systematically addressed. Clearly, using the fail-safe schools framework as a guide for improving school reliability would require creative use of school resources and difficult changes in school cultures. Nevertheless, as an effort to focus on organizational functions that are uncovered in the study of HROs, the fail-safe framework raises questions with intriguing possibilities.

## REFERENCES

- A safer runway? Can't tell. No accidents. (2002, October 27). *New York Times*, p. 39.
- Arif, M., & Smiley, F. (2003). Business-Like accountability in education: Commentary on Morrill to Baldrige. *Education*, 123(4), 740-762.
- Artiles, A., Harry, B., Reschly, D., & Chinn, P. (2002). Over-identification of students of color in special education: A critical overview. *Multicultural Perspectives*, 4(1), 3-10.
- Babb, J., & Ammons, R. (1996). BOP inmate transport: A high reliability organization. *Corrections Today*, 58(4), 108-110.
- Bain, W. A. (1999). Application of theory of action to safety management: Recasting the NAT/HRT debate. *Journal of Contingencies and Crisis Management*, 7, 129-140.

- Baldrige National Quality Program. (2002). *Educational criteria for performance excellence*. Gaithersburg, MD: Author.
- Beck, L. (1993). *Reclaiming educational administration as a caring profession*. New York: Teachers College Press.
- Beck, L., & Murphy, J. (1996). *The four imperatives of a successful school*. Thousand Oaks, CA: Corwin Press.
- Bierly, P., & Spender, J. (1995). Culture and high reliability organizations: The case of the nuclear submarine. *Journal of Management*, 21(4), 639-656.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.
- Blythe, T., Allen, D., & Powell, B. (1999). *Looking together at student work: A companion guide to assessing student learning*. New York: Teachers College Press.
- Burns, M., & Symington, T. (2002). A meta-analysis of prereferral intervention teams: Student and systemic outcomes. *Journal of School Psychology*, 40(5), 437-447.
- CalSTAT. (2003). *Award-winning district puts data to work*. Retrieved January 21, 2003, from <http://www.calstat.org/hesperia.article.pdf>
- Cameron, K. (1986). Effectiveness as paradox: Consensus and conflict in conceptions of organizational effectiveness. *Management Science*, 32(5), 539-553.
- Chalfant, J. C., & Pysh, M. V. (1989). Teacher assistance teams: Five descriptive studies on 96 teams. *Remedial and Special Education*, 10, 49-58.
- Coalition for Evidence-Based Policy. (2002). *Bringing evidence-driven progress to education: A recommended strategy for the U.S. Department of Education*. Washington, DC: Author.
- Cognition and Technology Group at Vanderbilt. (1997). *The Jasper project: Lessons in curriculum, instruction, assessment, and professional development*. Mahwah, NJ: Lawrence Erlbaum.
- Cohen, D., & Ball, D. (1998). *Instruction, capacity and improvement* (No. RR-42). Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.
- Crawford, L., & Tindal, G. (2002). Curriculum-based collaboration in secondary schools. In M. R. Shinn, H. M. Walker, & G. Stoner (Eds.), *Interventions for academic and behavior problems II: Preventive and remedial approaches* (pp. 825-852). Bethesda, MD: National Association of School Psychologists.
- Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment*, 7(4), 303-323.
- Darling-Hammond, L. (1999). Educating teachers for the next century: Rethinking practice and policy. In G. Griffin (Ed.), *The education of teachers* (Ninety second yearbook of the National Society for the Study of Education, pp. 221-256). Chicago: University of Chicago Press.
- DeFour, R. (2002). The learning-centered principal. *Educational Leadership*, 59(8), 12-15.
- DeFour, R., & Eaker, R. (1998). *Professional learning communities at work: Best practices for enhancing student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Deno, S. (1986). Formative evaluation of individual student programs: A new role for school psychologists. *School Psychology Review*, 15, 358-374.
- Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research*, 72, 433-479.
- Devaney, K., & Sykes, G. (1988). Making the case for professionalism. In A. Lieberman (Ed.), *Building a professional culture in schools* (pp. 3-22). New York: Teachers College Press.
- Doyle, W. (1983). Academic work. *Review of Educational Research*, 53, 159-199.

- Education Commission of the States. (2000). *Two paths to quality teaching: Implications for policymakers. An ECS staff comparison of the positions of the National Commission on Teaching and America's Future and the Thomas B. Fordham Foundation*. Denver, CO: Author. Retrieved April 21, 2003, from <http://www.ecs.org/clearinghouse/12/22/1222.htm>
- Elmore, R. (2000). *Building a new structure for school leadership*. Washington, DC: Albert Shanker Institute & American Federation of Teachers.
- Elmore, R., Peterson, P., & McCarthy, J. (1996). *Restructuring in the classroom: Teaching, learning, and school organization*. San Francisco: Jossey-Bass.
- Fiol, M., & O'Connor, E. (2003). Waking up! Mindfulness in the face of bandwagons. *Academy of Management Review*, 28(1), 54-70.
- Germann, G., & Tindal, G. (1985). An application of curriculum-based measurement: The use of direct and repeated measurement. *Exceptional Children*, 52, 244-265.
- Goldring, E., & Greenfield, W. (2002). Understanding the evolving concept of leadership in education: Roles, expectations, and dilemmas. In J. Murphy (Ed.), *The educational leadership challenge: Redefining leadership for the 21st century* (One hundred-first yearbook of the National Society for the Study of Education, pp. 1-19). Chicago: University of Chicago.
- Good, R. H., Simmons, D. C., & Kaméenui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading*, 5, 257-288.
- Goodlad, J., & McMannon, T. (Eds.). (1997). *The public purpose of education and schooling*. San Francisco: Jossey-Bass.
- Gresham, F. M., & MacMillan, D. L. (1997). Teachers as "tests": Differential validity of teacher judgments in identifying students at-risk for learning difficulties. *School Psychology Review*, 26, 47-61.
- Hall, G., & Hord, S. (1987). *Change in schools: Facilitating the process*. Albany: State University of New York Press.
- Hannan, M., & Freeman, J. (1984). Structural inertia and organizational change. *American Sociological Review*, 49, 149-164.
- Haynes, N., Emmons, C., & Woodruff, D. (1998). School development program effects: Linking implementation to outcomes. *Journal of Education for Students Placed at Risk*, 3(1), 71-85.
- Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of the literature. *Review of Educational Research*, 59, 297-313.
- House, E. (1998). *Schools for sale: Why free markets won't improve America's schools and what will*. New York: Teachers College Press.
- Huber Marshall, L. & Klebe, K. (2002). [Choice making for elementary students]. Unpublished raw data. Colorado Springs: University of Colorado at Colorado Springs, Center for Self-Determination.
- Ialongo, N. S., Edelson, G., & Kellam, S. G. (2001). A further look at the prognostic power of young children's reports of depressed mood and feelings. *Child Development*, 72(3), 736-747.
- Isenburg, M. (Producer). (2003, May 5). *Aurora images* [Television broadcast]. Aurora, CO: KACT-TV Aurora Community Television.
- Jesse, D., Davis, A., & Pokorny, N. (2004). Effective middle schools for Latino students in poverty. *Journal for Education of Students Placed at Risk*, 9(1), 23-45.
- Kaméenui, E. J. (1998). *Effective instructional support systems (EISS)*. Eugene: University of Oregon, Institute for the Development of Educational Achievement.
- King-Sears, M. E. (1999). Teacher and researcher co-design self-management content for an inclusive setting: Research training, intervention, and generalization effects on student

- performance. *Education and Training in Mental Retardation and Developmental Disabilities*, 34(2), 134-156.
- Langer, E. (1989). Minding matters. The consequences of mindfulness-mindlessness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, pp. 137-173). San Diego, CA: Academic Press.
- LaPorte, T. (1994). A strawman speaks up: Comments on the limits of safety. *Journal of Contingencies and Crisis Management*, 2, 207-211.
- LaPorte, T. (1996). High reliability organizations: Unlikely, demanding, and at risk. *Journal of Contingencies and Crisis Management*, 4(2), 60-71.
- Little, J. (1990). The persistence of privacy: Autonomy and initiative in teachers' professional relations. *Teachers College Record*, 91(4), 509-536.
- Louis, K., Marks, H., & Kruse, S. (1996). Teachers' professional community in restructuring schools. *American Educational Research Journal*, 33(4), 757-798.
- Lytle, S., & Cochran-Smith, M. (1994). Inquiry, knowledge, and practice. In S. Hollingsworth & H. Sockett (Eds.), *Teacher research and education reform* (Ninety-fourth yearbook of the National Society for the Study of Education, pp. 22-51). Chicago: University of Chicago Press.
- March, D. (2000). Educational leadership for the twenty-first century: Integrating three essential perspectives. In *Jossey-Bass reader on educational leadership* (pp. 126-145). San Francisco: Jossey-Bass.
- Marks, H., & Louis, K. (1999). Teacher empowerment and the capacity for organizational learning. *Educational Administration Quarterly*, 35(5), 751-781.
- Marston, D. (2003, February). Unpublished raw data presented at the meeting of the Pacific Coast Research Conference, La Jolla, CA.
- MacMillan, D., & Reschly, D. (1998). Overrepresentation of minority students: The case for greater specificity or reconsideration of the variables examined. *Journal of Special Education*, 32(1), 15-24.
- Meisels, S. J., DiPrima Bickel, D., Nicholson, J., Xue, Y., & Atkins-Burnett, S. (2001). Trusting teachers' judgments: A validity study of a curriculum-embedded performance assessment in kindergarten to grade 3. *American Educational Research Journal*, 38(1), 73-95.
- Meyer, J., & Rowan, B. (1978). The structure of educational organizations. In M. Meyer (Ed.), *Environments and organizations* (pp. 78-109). San Francisco: Jossey-Bass.
- Myers, V. M., & Kline, C. E. (2001/2002). Secondary school intervention assistance teams: Can they be effective [Electronic version]. *High School Journal*, 85(2), 33-43.
- National Staff Development Council. (1995). *Standards for staff development*. Oxford, OH: Author.
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).
- Nolet, V., & Tindal, G. (1994). Curriculum-based collaboration. *Focus on Exceptional Children*, 27(3), 1-12.
- Northwest Regional Educational Laboratory. (2002). *Catalog of school reform models*. Retrieved April 19, 2003, from <http://www.ncrel.org/scpd/catalog/index/html>
- Ortiz, A. A. (2001). Prevention of school failure and early intervention for English language learners. In A. J. Artiles & A. A. Ortiz (Eds.), *English language learners with special education needs* (pp. 31-48). McHenry, IL: Center for Applied Linguistics, Delta Systems.
- O'Shaughnessy, T., Lane, K., Gresham, F., & Beebe-Frankenberger, M. (2003). Children placed at risk for learning and behavioral difficulties. *Remedial and Special Education*, 24(1), 27-36.

- Oswald, D., & Coutinho, M. (1995). Identification and placement of students with serious emotional disturbance: Part I. Correlates of state child-count data. *Journal of Emotional & Behavioral Disorders*, 3(4), 224-230.
- Paris, S. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36(2), 89-102.
- Perrow, C. (1984). *Normal accidents: Living with high risk technologies*. New York: Basic Books.
- Perrow, C. (1999). Organizing to reduce the vulnerabilities of complexity. *Journal of Contingencies and Crisis Management*, 7, 150-155.
- Peterson, K., & Shinn, M. (2002). Severe discrepancy models: Which best explains school identification practices for learning disabilities? *School Psychology Review*, 31, 459-476.
- Pogrow, S. (1996, September 26). On scripting the classroom. *Education Week*, 16(4), pp. 52, 40.
- Reason, J. (2000). Human error: Models and management [Electronic version]. *British Medical Journal*, 320(7237), 1-6.
- Rijpma, J. A. (1997). Complexity, tight-coupling and reliability: Connecting normal accidents theory and high reliability theory. *Journal of Contingencies and Crisis Management*, 5, 15-23.
- Roberts, K. (1990). Managing high reliability organizations. *California Management Review*, 32, 101-113.
- Roberts, K., Bea, R., & Bartles, D. (2001). Must accidents happen? Lessons from high-reliability organizations. *Academy of Management Executive*, 15(3), 1-12.
- Roberts, K., & Libuster, C. (1993). From Bophal to banking: Organizational design can mitigate risk. *Organizational Dynamics*, 21, 15-26.
- Rochlin, G. (1993). Defining "high reliability" organizations in practice: A taxonomic prologue. In K. Roberts (Ed.), *New challenges to understanding organizations* (pp. 11-32). New York: Macmillan.
- Rochlin, G., LaPorte, T., & Roberts, K. (1987). The self-designing high reliability organization: Aircraft carrier flight operation at sea. *Naval War College Review*, 40, 76-90.
- Rosenholtz, S. (1989). *Teachers' workplace: The social organization of schools*. New York: Longman.
- Sammons, P. (1999). *School effectiveness: Coming of age in the twenty-first century*. Abington, UK: Swets & Zeitlinger.
- Secretary's Commission on Achieving Necessary Skills. (1991). *What work requires of schools*. Washington, DC: U.S. Department of Labor.
- Senge, P., Cambron-McCabe, N., Lucas, T., Smith, B., Dutton, J., & Kleiner, A. (2000). *Schools that learn: A fifth discipline fieldbook for educators, parents, and everyone who cares about education*. New York: Currency.
- Shavelson, R., & Towne, L. (2002). *Scientific research in education*. Washington, DC: National Academy Press.
- Shepard, L. (2003). The hazards of high-stakes testing. *Issues in Science and Technology*, 19(2), 53-59.
- Shimabukuro, S., Parker, M., Jenkins, A., & Edelen-Smith, P. (1999). The effects of self-monitoring of academic performance on students with learning disabilities and ADD/ADHS. *Education & Treatment of Children*, 22(4), 397-415.
- Shinn, M. R. (Ed.). (1998). *Advanced applications of curriculum-based measurement*. New York: Guilford.
- Skrla, L., & Scheurich, J. (2001). Displacing deficit thinking in school district leadership. *Education and Urban Society*, 33(3), 235-259.

- Slavin, R. E. (1996). Neverstreaming: Ending learning disabilities before they start. *Educational Leadership, 53*(5), 4-7.
- Slavin, R. E., & Madden, N. A. (1989). What works for students at risk: A research synthesis. *Educational Leadership, 46*(5), 4-13.
- Smylie, M., & Hart, A. (1999). School leadership for teacher learning and change: A human and social capital development perspective. In J. Murphy & K. Louis (Eds.), *Handbook of research on educational administration* (2nd ed., pp. 421-443). San Francisco: Jossey-Bass.
- Spillane, J., & Louis, R. (2002). School improvement processes and practices: Professional learning for building instructional capacity. In J. Murphy (Ed.), *The educational leadership challenge: Redefining leadership for the 21st century* (One hundred first yearbook of the National Society for the Study of Education, pp. 83-104). Chicago: University of Chicago Press.
- Stringfield, S., & Datnow, A. (2002). Systemic supports for schools serving students placed at risk. In S. Stringfield & D. Land (Eds.), *Educating at-risk students* (pp. 269-288). Chicago: National Society for the Study of Education.
- Sugai, G., & Horner, R. H. (2002). Introduction to the special series on positive behavior support in schools. *Journal of Emotional & Behavioral Disorders, 10*(3), 130-135.
- Talbert, J. (1993). Constructing a school-wide professional community: The negotiated order of a performing arts school. In J. Little & M. McLaughlin (Eds.), *Teachers' work: Individuals, colleagues, and contexts* (pp. 164-184). New York: Teachers College Press.
- Taylor, F. (1911). *The principles of scientific management*. New York: Norton.
- van Kraayenoord, C., & Paris, S. (1997). Australian students' self-appraisal of their work samples and academic progress. *Elementary School Journal, 97*(5), 523-537.
- Walker, H. M., Severson, H., Stiller, B., Williams, G., Haring, N., Shinn, M., et al. (1988). Systematic screening of pupils in the elementary age range at risk for behavior disorders: Development and trial testing a multiple gating model. *Remedial and Special Education, 9*(3), 8-14.
- Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly, 21*(1), 1-19.
- Weick, K. E. (1987). Organizational culture as a source of high reliability. *California Management Review, 29*, 112-127.
- Weick, K., & Sutcliffe, K. (2001). *Managing the unexpected: Assuring high performance in an age of complexity*. San Francisco: Jossey-Bass.
- Weick, K., Sutcliffe, K., & Obstfeld, D. (1999). Organizing for high reliability: Process of collective mindfulness. *Research in Organizational Behavior, 21*, 81-123.
- Wheatley, M., & Kellner-Rogers, M. (1995). Breathing life into organizations [Electronic version]. *Journal for Quality & Participation, 18*(4), 1-5.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64-70.

**G. Thomas Bellamy** is vice chancellor for academic affairs at the University of Washington–Bothell. His research interests include school leadership and improvement. His most recent publication is the chapter “A Model for Building Knowledge for Professional Practice,” coauthored with R. Muth, C. L. Fulmer, and M. J. Murphy, in *Educational Leadership: Knowing the Way, Going the Way, Showing the Way* (Scarecrow Education, 2004).

**Lindy Crawford** is an assistant professor in the College of Education at the University of Colorado at Colorado Springs. Her research interests include large-scale writing assessment, test

### 30 Educational Administration Quarterly

*accommodations and modifications for diverse learners, alternate assessments, and instruction for English-language learners. Her most recent publication is an in press article titled "Effects of a Read Aloud Modification on a Standardized Reading Test," coauthored with G. Tindal, in Exceptionality.*

***Laura Huber Marshall** is an instructor in the College of Education at the University of Colorado at Colorado Springs. Her research interests include student self-determination. Her most recent publication is a 2004 article titled "A 3-Year Study of Middle, Junior High, and High School IEP Meetings" in Exceptional Children.*

***Gail A. Coulter** is an administrator in the Montgomery County Public Schools in Maryland. Her research interests include reading interventions, reading process, and professional development.*