

## Evidence-Based Leadership by Design

Carol Barnes, Charles Vanover<sup>1</sup>, and Ji-Soo Kim<sup>2</sup>,

University of Michigan

Paper for the Symposium:

Scaling Up and Sustaining Educational Improvement Programs in

38 schools in 6 districts in 5 states

Chair and Discussant: Cynthia Coburn, University of California, Berkeley

Conference of the American Educational Research Association

Division L, Curriculum and Instruction, 2009

Research reported in this paper was conducted by the Consortium for Policy Research in Education (CPRE). The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of CPRE, the subjects studied or the sponsors of this research study.

**DRAFT: NOT FOR CITATION OR CIRCULATION WITHOUT PERMISSION  
OF THE AUTHORS.<sup>3</sup>**

---

<sup>1</sup> Equal, co-first authors

<sup>2</sup> Third author

<sup>3</sup> Please contact the corresponding authors: Charles Vanover [cvanover@umich.edu](mailto:cvanover@umich.edu) and Carol Barnes [barnscl@umch.edu](mailto:barnscl@umch.edu).

## ABSTRACT

Using data from semi-structured interviews and a school leader questionnaire from the Study of Instructional Improvement, we examine how leaders in two CSR designs—America’s Choice and Success For all—engage in the practice of evidence-based, instructional leadership. The goal of our analysis is to learn to what extent, and how, the school leaders learned to create sustained, school-wide systems of inquiry to collect, analyze, and use instructional evidence to improve professional practice and student learning.

## INTRODUCTION

We have seen at least a decade long call for a shift in school principals’ role from administrator toward instructional leadership in schools as one way to improve chronically underperforming **these** schools (e. g. Elmore, 2000; Fink & Resnick, 2001; Murphy, Elliott, Goldring, & Porter, 2007; Spillane, Diamond, & Jita, 2004). A key component of this kind of leadership includes using evidence based instructional guidance and decision-making. The problem has been how to “scale-up” this kind of leadership substantively, into many schools over time. We know of little evidence that this has occurred, especially in schools needing improvement.

In this paper we use two recent attempts to implement designed systems for improving leadership practice in high poverty schools. We examine how leaders in two CSR designs—America’s Choice and Success For All—engage in the practice of evidence-based, instructional leadership. In doing so, we focus on a set of critical interdependent resources for school improvement that make up a *sustained system of*

*evidence-based leadership practice* (Datnow & Stringfield, 2000; Elmore, Peterson, & McCarthy, 1996; Fink & Resnick, 2001; National Research Council, 2003; Paredes Scribner, 1999; Pellegrino, Chudowsky, & Glaser, 2001; Spillane & Sherer, 2004; Supovich & Klein, 2003). We investigate how leaders in a group of high poverty schools learned to gather information about teachers' teaching and students' learning and then use this evidence to improve professional practice and children's learning outcomes.

All of the schools we studied are located in districts engaging in standards' based reform. Under this theory of school improvement, every teacher in every school is required to provide each student with the learning opportunities necessary to complete an ambitious and cognitively-demanding curriculum (O'Day & Smith, 1993)<sup>4</sup>. No student can be allowed to fail. For the intentions of the reform to be realized, all students must engage in scholastic tasks that once were the province of a small elite (Resnick, 1995; Skrla & Scheurich, 2001). In standards-based educational systems, school leaders' primary responsibility is to ensure that every student in their building learns how to read, write, and skillfully participate in rigorous academic work. As many studies of high performing schools show, leaders cannot fulfill this responsibility unless they understand how well their teachers' teach and how successfully their students learn.

The schools we studied used CSR designs (Berends, 2004) as a means to achieve these ends. These externally designed educational reform programs provided leaders, teachers and other school professionals with what were, at the time of our study, new resources such as management routines, tools, and professional development services

---

<sup>4</sup> As many writers have pointed out (e. g. Allensworth, Corraera, & Ponisciak, 2008; Booher-Jennings, 2005; Diamond & Spillane, 2004), the enacted systems of teaching and learning that standards based school systems create may be quite at variance with these ideals. See Barnes, Massell and Vanover (forthcoming) for an analysis of the enacted educational policy environments in the group of schools we sample.

(Cohen & Ball, 1999; Cohen, Raudenbush, & Ball, 2003) for meeting the imposing goals set by standards based reform. We use the terms “*sustained or learned systems of evidence-based leadership practice*” to describe the interaction of these resources in the situated context of particular schools. This idea highlights the key role that continual learning plays in school improvement. To create high performing schools, teachers and school leaders must learn to diagnose their students’ strengths and weaknesses and then, over time, evaluate the effects of their choice of instructional strategies (Pellegrino et al., 2001; Resnick & Glennan, 2002; Scribner, 1999; Shepard, 2000; Skrla, Scheurich, Garcia, & Nolly, 2004). Regular, iterative feedback plays a key role in these learned systems of school improvement (Stringfield, 1995; Supovich & Klein, 2003). In this paper we focus primarily on the role of instructional leaders who must do more than learn how to collect and analyze assessment data. They must learn how to evaluate the validity of the evidence they collect, reflect on that information, and then use it mindfully to support instructional improvement. .

Constructing these “*sustained systems of evidence-based leadership practice*” at the local level thus poses complex problems of professional development, coordination, and performance (Coburn & Talbert, 2006; Murphy et al., 2007; Stein & Nelson, 2003). But the literature on planned organizational change in education demonstrates that school professionals do not have to invent all of this process of school improvement on their own (Berends, Bodilly, & Kirby, 2002; Camburn, Rowan, & Taylor, 2004; Crandall & Loucks, 1983; Datnow & Stringfield, 2000; Desimone, 2002; Firestone & Corbett, 1988; Rowan, Camburn, & Barnes, 2004; Supovich & Poglinco, 2001). Rather, leaders can use

the resources from externally designed educational reform programs, such as CSR designs, to reconfigure the way their schools and classrooms work.

To do so educators “must first notice, then frame, interpret, and construct meaning [from these resources]” (Spillane, Reiser, and Reimer, 2002; See also Cohen, Radenbush and Ball 2002). In other words educators must make sense of the elements of *evidence-based system* the designs provide, and then learn to use these resources, intelligently, for their students’ benefit. Externally designed programs of change do not implement themselves. They only add value, when they are activated by professionals working in schools. In what follows we review the literature that informed our conceptual frame, then explore the extent to which and how, the leaders in our study were able to enact an evidence-based practice based on externally designed, school improvement systems.

## **CONCEPTUAL FRAME**

### **School Organizations as Interpretive Systems**

Our paper is organized around a set of assumptions about how organizations function. We view organizations as interpretive systems that can process information in a manner analogous to the human mind (Daft & Weick, 1984; Hutchins, 1995a; Walsh, 1995; Walsh & Rivera Ungson, 1991; Weick & Roberts, 1993). Organizations think. They process information about what is happening inside and outside the system of activities that structure members’ work and use this knowledge to take action. Participants do not have perfect information; rather, in Daft and Weick (1984)’s words:

Organizations must make interpretations. Managers literally must wade into the ocean of events that surround the organization and actively try to

make sense of them. Organization participants physically act on these events, attending to some of them, ignoring most of them, and talking to other people to see what they are doing (Braybrook, 1964)(p. 286).

Leadership practice within these interpretive systems is a form of “information-work.” Leaders positional authority has the power to influence or “structure” the kinds of evidence organization members collect and to focus their attention on it (Spillane & Sherer, 2004). In schools, for instance, leaders might ask all primary teachers to take a running record of their students reading performance once a quarter. Managers also have the power to guide the flow of information across the different levels of their organization (Barnard, 1938). For instance, school leaders can to some extent alter the pattern of interactions among teachers, from working alone in isolated classrooms to regularly communicating in weekly meetings thus helping teachers share information about solving common problems. But, managers can also gather information that is never understood or used—mindfully or otherwise. Instead, as Feldman and March (1981) emphasize, inert information may exist primarily to give managers’ work the appearance of rationality. If information is to actually change how organizations operate and ‘think’, then members must notice it, talk about it, make sense of it and use the evidence they collect.

### **Routines: Ostensive and Performative Dimensions**

One of the major ways that organizations gather and use information is through routines (Dutton, Worline, Frost, & Lilius, 2006; Feldman & March, 1981; Feldman & Pentland, 2003; Feldman & Rafaeli, 2002; Leavitt & March, 1988; March & Simon, 1958). Routines are standard programs of action that organizational participants use to solve reoccurring problems. They are collections of tools and learned activities that

organization members apply to the challenges they face. Routines are viewed as effortful performances that may require skill to enact and learning to enact well.

Contemporary organizational theorists argue that routines have two dimensions (Becker, 2004; Feldman & Pentland, 2003). They exist both as ideas and performances. The *ostensive* aspect is the ideal or schematic form of a routine. It is the routine in principle. In this paper, we assume that some of the ostensive knowledge in the routines we studied was distributed in the tools that the two CSR designs gave organizational participants to use when performing recurring sequences of activities. We thus view, for instance, multiple forms of assessments, specified curricula, guides to action in manuals, or standards of performance as forms of ostensive knowledge (see e. g. Hutchins, 1995a, 1995b).

The *performative* dimension of a routine is the “situated action” constructed when participants use the routine in particular contexts. These performances are constructed from a common, learned grammar of action (Feldman, 2003; Feldman & Pentland, 2003). We found that routine performances varied in our CSR schools—across and within designs. Organizational members could perform a routine by rote, without much thinking or understanding, or they could perform a routine mindfully, consciously following the rule because they understood and valued it. Moreover, some routines tended to provide enactors with stronger exchanges of information and more opportunities for them to make sense of the content underlying or embedded in the routines.

### **The Designs for Instructional Improvement**

The America’s Choice (AC) and Success For All (SFA) CSR designs provided schools with powerful tools for meeting the challenges set by standards based reform.

Local leaders were not required to identify, collect and coordinate all of the curricular materials, aligned assessments, and organizational practices necessary to promote ambitious instruction. Nor were they required to organize and manage a system of evidence-based practice on their own. Instead, both the AC and SFA designs created a new leadership role: the CSR coach or facilitator. These newly created school leaders were expected to work with principals and teachers to improve classroom instruction. Moreover, the designs specified regular routines and tools for collecting evidence that were embedded into their systems of teaching or the leaderships' recurring transactions with teachers. These routines supported motivated leaders' efforts to monitor instruction and provide teachers with relevant feedback.

In this paper, we examine to what extent and how leaders in two CSR designs—AC and SFA—learned to create sustained, school-wide systems of instructional inquiry in order to collect, analyze and use instructional evidence for improving professional practice. Thus we ask:

- How do externally designed routines for evidence-based leadership shape leaders work in school implementing AC and SFA?
- How do these routines vary within and across CSR designs and schools?
- How, and to what extent do leaders learn to use these routines skillfully and mindfully to enact “sustained systems of evidence-based leadership practice?”

## **METHOD**

All data are from the Study of Instructional Improvement (SII), a longitudinal study of three nationally disseminated comprehensive school reform programs:



Accelerated School Project, AC, and SFA. In our paper we use primarily qualitative data: structured interview data with 38 principals implementing either SFA or AC in 2001; and semi-structured interview data collected in the spring of 2002 from a sample of school leaders—coaches, principals and vice principals in 8 AC or SFA case study schools participating in the larger quantitative study. Interviews were transcribed, coded based on a conceptual scheme, then entered into an NVIVO database for initial analysis. The authors further coded and analyzed the data for this paper using the analysis tools in NVIVO—for example, all mentions of leaders’ professional learning opportunities and their evidence-based leadership practice. Separate codes were created to index principals and coaches’ accounts of their work gathering instructional information, creating meaning out of these findings, and using this evidence to change professional practice.—

The 8 case study schools we studied in depth in the Spring 2002 interviews enrolled large proportions of minority and-or free lunch students. Sites ranged in size from Redmond Elementary in Florida, a school enrolling only 285 students, to Westwood Elementary, a school in the New York City system enrolling 650 (for more details on these schools see Table 1.0, Appendix A)

In order to describe school leaders’ evidence-based leadership practices and their professional development experience--beyond what our case study school or principal interview data could illustrate--we used the SII school leader survey data collected in the second year of the survey (spring, 2002). This wave of data corresponds to the second wave of case study interviews conducted that year. We focus in this paper on describing and comparing school leaders’ practices and professional learning—through using the designs in their schools or through external training in AC and SFA schools. We thus

exclude the Accelerated School Project leader sample. The data include comparison schools which did not adopt either of these programs. The final sample for the quantitative description in this paper is 301 school leaders in 86 schools (Table 2).

Table 2. Quantitative Samples (N=301 school leaders in 86 schools)

| Program    | Leaders |       | Schools |       |
|------------|---------|-------|---------|-------|
|            | N       | %     | n       | %     |
| AC         | 130     | 43.2  | 31      | 36.0  |
| SFA        | 89      | 29.6  | 29      | 33.7  |
| Comparison | 82      | 27.2  | 26      | 30.2  |
| Total      | 301     | 100.0 | 86      | 100.0 |

We created five factor score scales in order to describe leaders' perceptions of their evidence-based leadership practices and their professional development experiences. We created two measures to describe leaders' professional development experiences during the school year. (1) PDFOCUS indicates the extent to which school leaders report that they have received professional development focusing on evidence-based leadership practice. (2) PLO indicates the degree to which leaders perceive that their professional development was effective and useful for their practice. The measures we use to describe leadership practices include: (3) WORKTCH, how often leaders observe; share information or advice; examine student work, and standardized norm-referenced test results; and work directly with teachers. (4) MONITOR indicates the frequency of leaders' engagement in various monitoring activities related to the improvement efforts.

And, (5) INFO-USE measures the extent to which school leaders' report using various sources of information.

Using these factor score scales, we explored the differences between AC, SFA, and Comparison school leaders. We ran One-way Analysis of Variance (ANOVA) and post-hoc tests for multiple mean comparisons between programs. We also explored item-level difference in selected scales. Additionally, we create school mean scores per scale by aggregating individual leaders' scores within a school. To locate our case study schools in the larger survey data sample and to identify the extreme cases, we then ranked the schools in order.<sup>5</sup>

## FINDINGS

### **Success For All and America's Choice as Learned Systems of Evidence-based Leadership Practice**

We begin with a holistic account of what the leaders in our study said they were trying to do as they enacted the SFA or AC designs to ground our descriptions of leaders' practice in an inclusive framework. We thus describe the intent or 'ostensive' aspect of the designs, as well as enactors' implementation broadly, to contrast the evidence based system in each design. In subsequent sections, we discuss how school leaders learned to use the resources the two designs provided and how the designs shaped leaders' work in our schools. Finally we describe the *what* and the *how* (Spillane, Diamond et al., 2004) of evidence-based practice in greater detail to illustrate the obstacles leaders faced, and the challenges they were able to overcome. Throughout, we contrast SFA's more specified

---

<sup>5</sup> For a description of each measure, and more detail on both quantitative and qualitative methods and data see Appendix B.

evidence-based system of leadership to the less definitive, but highly active, systems of evidence-based practice in AC schools.

**Success For All** An evidence-based system was designed into the operation of the SFA reform model (Camburn et al., 2004; Peurach, 2005; Rowan, Camburn, & Barnes, 2004) at the time of our study. At the heart of the design was the practice of grouping students for reading by ability. In our SFA schools, the teaching of reading was organized around a specified curriculum of lessons organized like a ladder. Students stepped onto this structure in Kindergarten when their pre-reading skills were tested and then each child was placed into a specific level of SFA's Early Learning<sup>6</sup> program. Students then progressed upwards one structured lesson at a time skipping or repeating units, at least in theory, according to their individual growth and development.

The design specified many aspects of leaders' roles within this evidence-based system. At the time we collected our data, the SFA Facilitators' role was to support teachers in learning to use the designs instructional routines and to manage the schools reading program. Their goal was to use evidence to improve instruction and eliminate early-reading failure by ensuring that every student in the school read at grade level or above. The principals' job was to motivate school personnel to implement the program; that is, to use their authority to monitor and hold teachers accountable for faithfully implementing SFA instructional practices, as well as to create a school-wide vision for school improvement. For the most part, in the schools we observed we found that leadership was distributed (Camburn, Rowan, & Taylor, 2003; Spillane, Halverson, &

---

<sup>6</sup> For clarity, all CSR related programs, roles, and components will be capitalized.

Diamond, 1999) between the instructional expertise and program knowledge held by SFA Facilitators and the positional authority of the principal.

Generally, the design's formal and informal professional learning opportunities,<sup>7</sup> and the specified elements of the program allowed motivated leaders to use the 'ostensive' knowledge distributed in SFA resources. In the schools we studied, this knowledge was sufficient to implement much of the design rapidly. In the words of one of the SFA principals we interviewed:

Everything is real tight in SFA. We know what our staff development is going to look like. We know what the program is going to look like. We know how we're going to monitor the program. SFA is the only comprehensive school reform design that if you buy into their program, the truck backs up, you unload the stuff, you get your training, and you're off and running. [S2002]

As we will discuss, however, one area of difficulty that created within design variation in implementation were program features that were not specified clearly. Another was SFA principals' weak knowledge of the instructional design. In the schools we studied, facilitators rather than principals were the primary source of instructional leadership.

**America's Choice** Where the stated goal of SFA, at the time of our study, was to ensure that every student in every elementary school read at grade level, the objective of the AC comprehensive school reform design was to ensure that every student mastered

---

<sup>7</sup> These might include trainings, conferences, and site visits during which the external Facilitator worked with the leadership team. They might also include the SFA leadership manual—especially the instructions for enacting the design's quarterly assessment system and monitoring teachers' practice. Finally, leaders might also use teachers' guides to understand, monitor and guide classroom practice

the skills necessary to succeed in college. The design was based on an effort-based (Resnick, 1995) view of intelligence and human capacity. All students were held to be capable of learning complex and cognitively demanding academic skills.

The heart of AC's evidence-based system was a set of performance standards and formative, classroom assessments for judging student work based on those standards. The New Standards Performance Standards specified the academic tasks students were expected to complete at each grade level. By design, teachers were asked to monitor and guide students' progress in mastering these tasks. To accomplish these goals, AC provided schools with formative assessments such as 'Running Records' of reading fluency, as well as routine, but authentic forms of classroom assessments for students' writing. Teachers in our AC case sites learned to record their one-on-one conferences with students in 'Conference Logs' and other kinds of documents. Both teachers and students in the cases engaged in whole class activities such as 'Author's Chair' during which students read work they wrote for peer and teacher critique—again based on commonly held criteria embedded in the New Performance Standards.

Similar to SFA, the AC design specified many aspects of leaders' roles and organized their work around supporting and monitoring a complex system of evidence-based practice. At the time we collected our data, AC expected the principal to act as a strong instructional leader who understood all the aspects of the design and worked with teachers to improve students' learning outcomes. The school's 'Literacy Coordinator' was expected to work directly with classroom teachers in intensive coaching sessions focused on how to perform Writers Workshop routines to ensure that student would meet the design's writing standards. In our cases, a coach might spend six weeks or more in an

individual teachers' classroom. Moreover, these coaches teamed with principals to implement the evidence-based system through "focused walk" routines during which they would not only move through classrooms to check student work displayed on bulletin boards, but would also examine students' writing in the AC design's 'Source Books', and check on teachers' Conference Logs. At the time of our study, AC specified the role of a 'Design Coach who was expected to guide the implementation of the design while the Literacy Coach worked with individual teachers. But case principals varied dramatically in how they used this role, and the person filling it in two of our cases understood very little about the AC design.

### **Leaders' Professional Learning, Knowledge Use and Change**

Both designs encouraged school leaders to learn how to organize systems of evidence-based practice. These systems required leaders to perform new roles that could alter traditional work norms, use new routines for guiding instruction and engage with novel assessments or other tools. Leaders were given a wide variety of professional development experiences that included monthly trainings, yearly conferences, site visits from design liaisons, and written manuals for guiding practice. Generally, all of the school leaders we interviewed reported that they attended their respective design's conferences as well as participated in networks where they received training. These experiences allowed them to draw on the experience of colleagues with common, practical concerns related to design implementation. Principals and/or coaches in a majority of the schools also said they used the designs materials, and received technical assistance from the designs' liaisons who would visit schools to monitor and support implementation.

But we found distinct within and across design differences in the nature of leaders' reports about their learning experiences, engagement with the designs' resources, and thus, their understanding of the design—AC leaders reported they had more to elaborate and thus took longer to learn to use the designs. SFA leaders, on the other hand, talked less about learning, and more often about the specified nature of the SFA system, or the “technique” they and the teachers had acquired. The initial implementation of the SFA design also seemed to take less time than did the AC design. In general, coaches from both designs reported stronger learning experiences than did principals.

### **Principals**

**Success For All.** A key theme among the 8 SFA principals we interviewed in 2001 was a lack of engagement in learning about or otherwise using the design; that is, a distance from the design, especially its instructional features. Three of these principals never mentioned working with a network, for example, and only 3 of the 8 principals mentioned receiving technical assistance from SFA liaisons. Instead, the school's SFA Facilitator was the primary contact with the design and the person who reported involvement in the design's networking and technical assistance activities.

We found that only 2 SFA principals reported changing their practice to a large extent because of the design. Importantly, when principals did report changing their practice, the nature of that change was not substantial, and it limited the kind of information they could gather. For example, when asked directly if his role in the school changed as a result of SFA, one principal said, “I wouldn't say it has changed per se. I'm still basically doing the things I was before. I'm still in the classroom like I was in the classroom [before SFA]”. Another talked about resources, “I now have a facilitator for



reading and I work with them in coordinating the goals and objectives and needs of the program. That gave me some additional support staff. . . . I don't think that my role has changed that much.” Consistent with the “accountability” role the design emphasized one SFA principal said: “I had to make a commitment to the fidelity of the program and guard against any tampering with it”. Still another said: “Our school did not have the space to accommodate an SFA program, so I spent time making space, so we could follow the SFA guidelines”. None of these reports demonstrate knowledge of instruction or much interest in the content underlying the SFA instructional routines. In our analysis, most of the SFA principals’ we interviewed demonstrated quite weak knowledge of the design. This is a critical issue because to “monitor” classrooms effectively, principals would need to have some understanding of the intended content in the instructional routines teachers used (e. g. Murphy et al., 2007; Stein & Nelson, 2003). Without this kind of knowledge they could not be effective at providing guiding feedback, or even at overseeing implementation.

During our Spring 2002 interviews, we asked for more information about principals’ work as instructional leaders. Most of the SFA principals-in our case sites reverted to traditional classroom observations that typically lasted five minutes or less. Three told us they used what one described as a “cheat sheet”—a schedule of the activities teachers were expected to perform during specific points during the lesson—to enforce pacing and content coverage. None of these same SFA principals told us they monitored the quality of teacher-student interactions or student performances. Nor did they report engaging much in learning about the instructional designs through the monitoring routines. Instead, they observed instruction mostly by rote. The lack of

understanding made the work of the facilitators—most of whom did have a strong understanding of the instructional designs— more difficult and, according to these facilitators, sometimes delayed teachers progress in moving beyond a more superficial enactment of the design. One of the reading facilitators at Doris Elementary told us when we asked her to describe the obstacles she faced doing her job:

The problem I have is that the administration were not properly trained in the program. More training is needed for them as supervisors, of what they need to monitor, so that they know as they're walking through a room. [..]First of all you have to know what the program is about. You have to know all the pieces. You have to know over time where people have been. It's not fair to walk in on someone one day and not come back for a week, three weeks, not come back to see what those kids are doing. It's sort of like, throughout the 90 minutes [the principal should] be an active part in the program. (S2002)

Thus, by this account, not only did the principal not observe instruction frequently, she did not understand enough about it to be effective. Principal Shur at Doris explicitly confirmed this report saying, “I don't know SFA that well.” Shur told us she never found the time during the day to attend an SFA component meeting. As a result, in her account, she was not able to take advantage of an important learning opportunity.

I have to be honest with you. As much as I want to and I make commitments to my Facilitator, I said, “Please let me know, I want to come to the meeting.” And then, as administration, I get so involved in the lunch periods or the paperwork that I can't get to it. And that really does

bother me, because I know I need to be there, and I need to let the teachers show that this is very important to me. I've made a time to come, and every time I've tried to come, something's come up. A parent's walked in, a child's in crisis. (S2002)

Immediately after this part of the interview, Shur told us that she held one of the most important jobs in the SFA design--Roots Supervisor--but admitted that she did not have enough knowledge to manage this aspect of the design effectively.

In this SFA school, the principal's lack of understanding, commitment or inability to put aside other pressing administrative tasks, weakened both the SFA resources for monitoring the quality of teachers' implementation and the design's incentives for surpassing more than a mechanical use of routines. The design's elements for gathering instructional information and its elements for using it to improve instruction were not mutually reinforcing. Instead, the routines operated in a fragmented fashion that did not support an iterative improvement process. While Doris Elementary was the most definitive case and our weakest CSR implementation, this was the trend across all the SFA case study schools we investigated in 2002. We also found accounts that suggested a clear lack of knowledge of the design in 6 of 8 principal interviews we collected in 2001.

*America's Choice*. In contrast, AC principals frequently used the design resources—new routines for changing time allocations and gathering data for example, or curricula and rubrics specifying common student benchmarks--to become more engaged instructional leaders. For example, most of the AC principals we spoke to reported using the New Standards Performance Standards to assess students' written work during grade level meetings, classroom observations or in other forums. Twelve of the 13 AC

principals we interviewed in 2001 reported not only changing their role in the school, but also learning from using the information the AC's system of routines and tools generated.

For example, one AC principal said:

I look at the classroom differently. I talk with the students more. I look at students' work more. I scrutinize the lesson plans more. I look to see if the lessons are geared towards the standards. I watch the way teachers respond to students' questions. I look for different kinds of writing materials and read the student writing samples. I watch the way teachers ask comprehension questions of the students. I tune into the responses of the students. I work differently with the teachers. [AC S2001].

This comment was qualitatively different than any of the SFA principals' responses to the change question—even the stronger principals. It focuses primarily on how the principal was learning to assess instructional interactions and a range of student performances based on commonly held criteria—performance standards. Only 1 of the SFA principals we interviewed described understanding classrooms well enough to “tune into the responses of the students.” To do this work, many AC principals also reported the design's professional development and instructional leadership routines had quite dramatically increased their time in classrooms. When asked, for example, if AC had changed her role in the school an AC principal said:

[I am] in classrooms 2.5 to 3 hours everyday. More tuned into the instruction that's happening in the school. I like the fact that I can say, "Do not bother me, as I'll be in the classrooms -- only contact me if the school's burning". This took

some re-learning from my office staff who were used to having me around all day. [S2001]

This account contrasts with SFA principals who more often reported spending a few minutes in classrooms they observed. Notice also, that the principal and her staff were learning and changing as they used the design. This was also the case for the AC principal just below who describes a newly focused role, even as she was still “figuring out” how to enact it:

[My role is] more difficult because now the most important thing I do is listen to children, read with children, look at their work. When I talk with teachers, I now talk about their instruction. . . . I can't figure out which of my other duties to give up either. Plus learning what to do with all of the assessments we're doing. How to really make that effective for instructional use? How to be able to give student test results to teachers in a manageable way so they can quickly relate it to instruction? [S2001]

We heard about this kind of learning on the job—in and from practice as leaders attempted to enact AC routines—from principals across most case sites. AC principals reported being engaged in instructional improvement and were knowledgeable about the details of their CSR program. However, the principals also told us that making sense of their new leadership routines and putting it all together into a working system took time. In part this may have been due to the AC design that intentionally left some aspects of evidence-based leadership routines open for elaboration by leaders in the context of their schools. Eight of the 13 principals we interviewed in 2001 for example, reported

difficulties in figuring out exactly what the design wanted their schools to do. None of the SFA principals we interviewed at this time mentioned such a compliant.

### **Facilitators or Coaches**

**Success For All.** Most SFA Facilitators worked as reading coordinators or teachers before their school adopted SFA. Unlike their principals, facilitators told us that the design’s resources improved their capacity to benefit their students in very concrete ways. All of the SFA Facilitators we interviewed in 2002 specifically mentioned learning from SFA trainings. Three of them provided descriptions of acquiring “useable” tools—new and well-defined “techniques”—that were especially useful to their own practices as instructional leaders. For example, when asked about her learning, the SFA Reading Facilitator at Nightingale elementary, reported, “I’ve learned reading strategies and techniques that I never knew existed.” She continued to explain these practice based “techniques” were missing from most formal pre-service education for teachers including her own:

We never had anything about teaching kids how to read or how to do math [in college]. And even. . . the kids coming out of college now—because I had student teachers—didn’t have those skills. I know the colleges frown on SFA because it’s too structured and it doesn’t allow creativity for their. . . student teachers. . . [But] I feel that a lot of those [strategies] have helped us. I mean, our test scores have gone up, our reading levels have gone up over the years—the number of kids reading at or on level. . . I think our school was lacking direction and structure. [S2002]

Similar to reports from Nightingale, the SFA Facilitator at Grapple Elementary reported success with using the SFA “structures” and “training for best practice”:

When we started SFA 11% of our kids were reading at grade level. Now we're up to over 50%. So obviously there was a need there that wasn't being met. They [SFA] provide the structure. The actual training of the best practices for what will work. [S2002].

These comments and others across our cases show that SFA resources provided leaders and teachers with especially useful knowledge—not only to inform Facilitators' understanding of instruction, but also ‘ready-made’ to implement in classrooms. Moreover, the ‘ostensive’ aspects of the design—scripted routines and content for example—were specified enough to structure educators' attention on classroom practices that could improve early reading instruction (e. g. Rowan, Camburn, & Correnti, 2004; Rowan & Correnti, 2007). Leaders in most SFA case schools learned to ‘perform’ classroom routines with fidelity and then support teachers in doing the same, in the context of their schools. Moreover, most case study Facilitators spoke very positively about this learning process in terms of implementing the design for reading in the lower grades. The Facilitator at Nightingale Elementary, for example, explained how learning from SFA conferences supported the school leadership team's efforts to take action based on its assessment that teachers were not effectively teaching the more complex, upper elementary SFA reading component.

I learned a lot at that conference. And I came back and we had a workshop the following week and I was given permission [from the principal] to take all the [SFA component] teachers aside and take them through Wings step

by step from the beginning and re-teach, retrain them how to do Wings, because I learned it. And that was the most beneficial [conference] I ever went to, because we all learned all over again how to teach Wings.

[S2002]

Here, in line with the themes we have discussed in this section, the Facilitator described how design based knowledge, and the authority to put this knowledge to work, in her view, supported the school's efforts to improve the implementation of SFA.

*America's Choice*. Recall that the nature of the AC design was generally less specified for both teachers and school leaders. The design provided a blueprint for action that linked classroom cycles of assessment to school wide monitoring of student progress—but both teachers and school leaders had more to “invent” or elaborate than did SFA leaders. This was especially so for AC's Readers Workshop components. Across our interviews, AC leaders and teachers reported that while Writers' Workshop provided concrete guidance for implementation, Readers' Workshop did not. More than one AC Literacy Coordinator described the difficulty of learning to teach teachers from a loosely specified reading design. For example:

LC. Sele: Well the writing was the most successful maybe it was the easiest to do [...] and the Guided Reading was the most difficult.

I: Was it? And why?

LC. Sele: Because I don't think the teachers understood what Guided Reading was [...] You have to really get the feel for [Guided Reading] and understand what it is and know what the kids need. It's almost something that you can't teach. As much as I could go in and model, I'm doing it



with a teacher now: She can't get her class managed so that she can watch what I'm doing. She can't manage. She can't get the kids in centers and we've done it on a repeated basis. She can't manage the class. If you can't manage the class you can't do Guided Reading. [S2002]

At this point in the implementation process—4 years after the school first adopted AC—Ms. Sele the Literacy Coach raved about the guidance she received from the Writer's Workshop component of AC. She believed AC's training strongly supported her efforts to teach this aspect of the design to the school's teachers. Among other elements, AC's professional development for leaders included concrete samples of student assignments called “core assignments keyed to the New Standards Performance Standards and clear guides for helping teachers construct classroom routines that embedded a formative assessment cycle into their writing instruction. This was not the case for Readers' Workshop as Ms. Sele understood it. For that program component, not only did she report having to invent materials for educating teachers, but she also felt the need for more direction about how to put the reading program into place, before she could adequately serve as a teacher educator.

Nevertheless, all of the AC Coaches reported they had increased their knowledge of instruction through the learning opportunities provided by the design—especially in writing. In the spring of '02 Ms. Sele explained:

Well I was trained almost four times a year by America's Choice...for a full week and they were fabulous opportunities for me to learn and that was the only way I could've come back and taught the teachers to do it...They were fabulous.

When asked who taught her the AC methods she also had very positive descriptions about the expertise of her teachers. She said:

They were America's Choice facilitators--fabulous teacher/trainers that came and trained us. They had experts come in. . . . and teach us writing. They were just fabulous people. So we felt we had the best training and that's what really made me able to come back and do it. (Ms. Sele, Spring '02)

When the interviewer asked, "And did you have materials? Did they give you materials?"

Ms. Sele responded:

They gave us something they call core assignments [in writing] so the teachers would know what steps to take in a narrative account and they guided us to what books to buy for the school. And when we had the funds, we bought every book that we possibly could buy to support the America's Choice training...The training was worth every penny that we paid for it and more. (Ms. Sele, Spring '02)

Ms. Sele's positive comments about AC training are echoed by school leaders at Bonds Elementary. Leaders at this AC school not only emphasized that the model's explicit learning opportunities had greatly improved their ability to lead instruction, but they also commented on the complex nature of the change the design required. Ms. Calm, the school's Design Coach, told us that the school never formally taught writing prior to implementing the AC Design and that leading others in the subject was new to her. When asked if she found the AC trainings helpful and if she was able to use what she learned in her practice, Ms. Calm said:

Oh sure. I have learned a lot. . . . Teachers in America’s Choice have to be able to observe what’s needed and know what to do next. They have to put all the connections together to get the end product that we all want... [The America’s Choice Design] is an awful lot of work, but it provides a mechanism for which you [teachers] can operate within to help your children be the best that they can be. Never before have we at Bonds formally taught writing. It was a mishap of reading. Before [AC], reading and writing was “Reading something and answering the questions.” That was basically the extent. [We] weren’t doing the high-level thinking on a large scale. So, America’s Choice, it sort of makes you stretch. (S2002)

This type of change was not reported for Readers Workshop. Teachers and leaders at Bonds, similar to those in Westwood, reported they needed more concrete guidance for implementing the Readers Workshop routines.

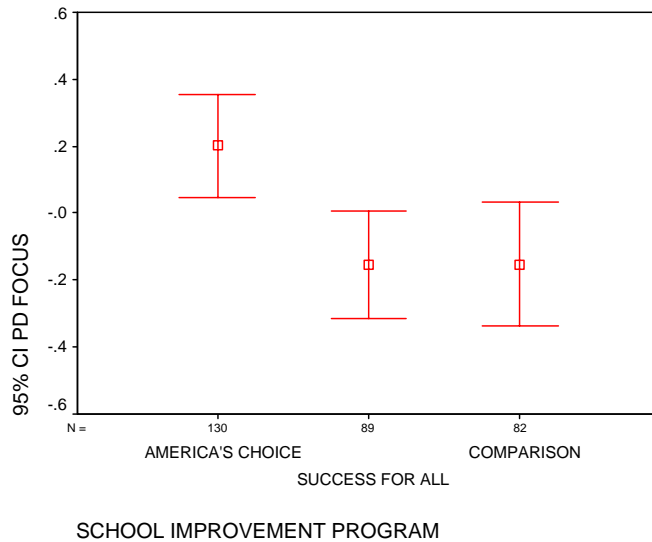
One point to notice in these accounts is that constructing high quality ‘performances’ from the instructional design is not only “an awful lot of work,” but requires that teachers and thus leaders elaborate on or “fill in” the routines they learn, using professional judgment that was not always supported by the design. Learning how to productively observe students’ performances, connect those performances to one’s teaching, decide “what to do next” and then elaborate on the design’s routines while constructing an instructional discourse in classrooms was complex. Thus, the nature of the AC design created special demands for leaders. But leaders across our case sites reported learning to perform AC “data gathering routines,” as well as working intensively in classrooms to engage teachers in a cycle of formative assessment, using a wide range

of information, that then fed back not only to instruction, but school-wide, to the leadership.

### **Quantitative Description of School Leaders' PD Experiences**

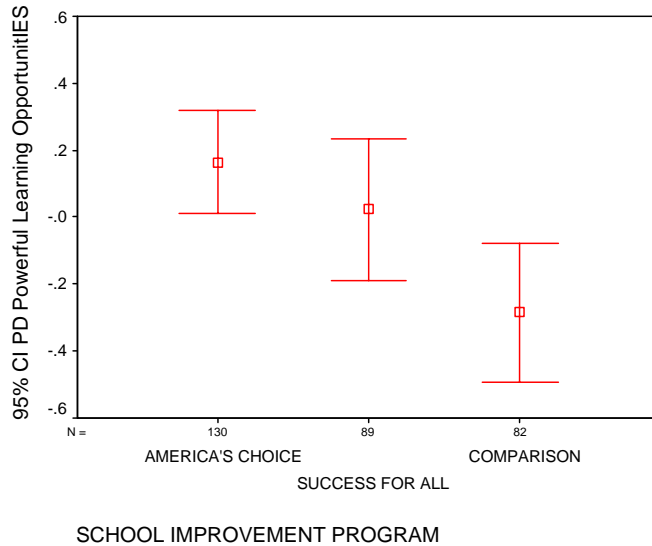
Generally, our quantitative descriptions of school leaders' professional development (PD) experiences complement the qualitative descriptions. Not surprisingly, Figure 1 below shows the AC leaders' mean score is higher than leaders in SFA and comparison schools for PDFOCUS. The mean difference between AC leaders and SFA leaders is statistically significant ( $p < .01$ ). But the means are not significantly different between SFA leaders and comparison school leaders. The items in this measure capture learning experiences related to some of the same elements of the "evidence-based system" that showed up in our AC principals' and coaches qualitative change reports. Leaders were asked to report whether their professional development included elements such as 'standards based learning', 'working with students', 'closely observing classroom practice', and 'new procedures for assessing students' performances.' These form the kind of evidence-based practice the AC design encourages its school leaders to us.

#### **Figure 1. PDFOCUS by Program**



We also found that compared to their counterparts in comparison schools, AC leaders more often agreed that their professional development was useful for practice as well as coherent and sustained PLO (see Figure 2 below). The mean difference is not significant between AC and SFA leaders. While the average score of SFA leaders is higher than that of comparison school leaders, the mean difference is only marginally significant ( $p=08$ , Table 1.3 in Appendix A). Still, together with our qualitative data, these results suggest that school leaders across both designs—those who were motivated and thus engaged with the designs’ resources—found their learning experiences useful for practice (based on the literature on adult learning, e.g., sustained, coherent, practice based development opportunities are more useful for change, than other kinds of professional development).

Figure 2 PLO by Program



### Systems of Evidence-based Leadership Practice in Action

In this section, we elaborate on how leaders interpreted and ‘performed’ particular routines in the two designs’ evidence based systems by drawing on their descriptions of daily practice. First, we show how specific design content and assessment routines shaped leadership practice differently across the SFA and AC schools. Our survey and case findings show that, from our framing perspectives, the ‘ostensive’ aspects of each design did indeed produce different systems of evidence-based practice. They also illustrate how leadership was distributed (Spillane et al., 1999) over different combinations of people, artifacts, and recurring situations in the schools we studied. These, in addition to the earlier reports of change and learning, illustrate the *what* (content focus) and the *how* (recurring actions) of evidence-based leadership practice using extant cases, and attempt to account for some of the variation across designs.

This final section also elaborates on a second theme that has run through this paper: performance of these situated systems of evidence-based leadership varied greatly *within* designs. The programs’ ostensive knowledge, the school context, and the capacity

and commitment of leaders all influenced leaders' ability to learn how to put the pieces of the design together into a functioning system and coordinate activities within that social structure. We discuss the difficulties leaders faced and show how some knowledgeable and committed leaders were able to overcome these challenges over time.

*Putting The Pieces Together: Gathering, Using And Making Sense Of Instructional Evidence*

We found that that leaders in our schools across designs collected both qualitative and quantitative evidence on teachers' teaching and students' learning. Nevertheless school leaders' implementation of data collection routines varied a great deal within designs, especially when leaders had to elaborate, select or invent tools for quantitative data collection. The extent to which leaders understood the assessments they used and could interpret (Daft & Weick, 1984; Paredes Scribner, 1999; Pellegrino et al., 2001) the data they collected also varied. While all leaders said they observed teachers, the designs as well as the features the enactors emphasized, shaped leaders' practices in qualitatively different ways. Thus, in practice, SFA leaders primarily focused on standardized assessment data and faithful implementation of SFA's instructional routines or scripts, while AC leaders focused on a broader range of qualitative evidence representing student performances and teachers' ongoing assessments of those performances.

**Principals and Facilitators in Success For All.** Our interviews with 7 of the SFA principals and all of the facilitators showed a general focus on the fidelity of implementation of the design's components. Reports of classroom observations, as a source of information in our SFA cases, ranged from formal evaluations—mostly conducted by principals or vice principals--to informal “pop ins” for checking the pace at

which teachers were covering material. In some schools teachers were required to post a report on their door, identifying where they were in the five-day SFA pacing schedule for the principal or SFA facilitator. When asked, for example, a Facilitator at Gladrock elementary reported observing and providing feedback to *most* teachers in the school each week: “Do I observe them [teachers] and give them feedback?. . . Yes, on a continuous basis.” When asked how long her classroom observations lasted she explained: “It can be anywhere from two minutes, just popping in, seeing what's going on. And it could be all the way up to five or ten minutes, depends”(F02).

The strongest SFA facilitator, Ms. Fina at Nightingale Elementary describes the benefits of these observations for gathering information about the status of SFA instruction:

Right now, I try to pinpoint certain areas of the teachers’ lessons. I have a graph on my computer that I used to keep track of what part of the lesson I’ve seen and when I saw it last. I can pull it up and say, “I haven’t seen this teacher do [SFA component] Adventures in Writing all year. Let me go in and see that now.” (S2002)

These very brief visits were common across SFA schools, but their effectiveness seemed to vary across roles. As we noted earlier, while we found that SFA facilitators generally understood the instructional program, and could productively provide teachers with the feedback they might need to guide instruction, SFA principals rarely did. This pattern of practice was helpful in keeping teachers on a fast paced schedule of content coverage—a pace that provided students with multiple opportunities to learn—but it was less conducive to using the observation



evidence to redirect teachers toward a less mechanical implementation of the instructional routines—especially in the upper grades.

We found that for the evidence based systems in the designs to work well in practice, its elements needed to be well coordinated and mutually reinforcing. As noted throughout, while principals engaged in “formal evaluations” and carried out the monitoring function in practice (see e.g. Camburn, Rowan & Taylor, 2004), they were often not well equipped to do so as they lacked knowledge underlying the SFA instructional precepts. Instead, for information about the quality of SFA instruction, principals relied on the SFA facilitators who were “teacher-leaders” as opposed to “administrators” and thus could not officially evaluate their peers. Though these Facilitators had a deep understanding of the SFA instructional system, all four we interviewed also told us they were very uncomfortable “evaluating” their colleagues due to traditional norms of teacher evaluation. When asked if she monitored instruction for example, the facilitator at Doris, our weakest SFA school said:

No. My personality is to be more of a people pleaser, to be perfectly honest with you. I have a hard time telling people that what they’re doing isn’t good. It’s an administrative thing and I don’t want to be an administrator.

Instead, facilitators worked to build a trusting relationship with their colleagues and for the most part were respected for their knowledge of the instructional design. Thus positional authority and authority derived from knowledge of the SFA program were distributed across leaders in our SFA schools but not necessarily mutually re-enforcing. Facilitators often felt constrained in critiquing their

colleagues and disliked informing principals about problems, but principals were unable to identify classroom problems on their own.

While the leadership functions, as implemented, varied across the principal and Facilitator roles, SFA leaders' use of quantitative evidence, clearly varied within roles. The SFA principals we interviewed in 2002 were once again distant from this aspect of the design, but organizing and managing school wide assessments posed many difficulties for all of the SFA facilitators we interviewed. This was especially the case for the upper elementary grades where, at the time of our study, the SFA design for Wings required school leaders to choose from a long list of possible assessments that ranged from computerized assessments to chapter tests from different Basal reading series. Many Facilitators reported they had problems finding a Wings assessment that fit the needs of their building.

Thus, implementation of this element of the design varied across our schools. The reading facilitator at Grapple for example, told us she tried and then abandoned at least four different tests for the design's 8-week assessments. Ms. Zann at Doris elementary school had similar problems and told us flatly that Wings had no diagnostic assessments, "no matter what the [SFA] Foundation says." While, Ms. Zann started using the Gates McGinnity quarterly assessments, she found them to be biased and changed to SRI's computerized tests. She received no training for using this feature of the designs; rather, by her account Ms. Zann took the schools' SRI data home every 8 weeks and had her husband chart it out and analyze it on his computer.

In contrast, the Facilitator at Nightingale, our strongest case of leadership, did receive extensive training in how to read the SFA quarterly assessment results, though

not until three years into the implementation of SFA at her school. Still, Ms. Fina told us that this training greatly improved her ability to lead instruction. She explained:

I never knew anything about data. I was actually one of those people, who was afraid to share data because it never looked really good and I was like, “I don’t want to share this!” But I’ve learned. I’ve learned how to look at it, and I’ve learned how to. . . say, “Well, this is why this isn’t working. This is why these groups of students aren’t moving.” (S2002)

Thus, the extent to which SFA school leaders, even facilitators were able to use the design elements in the evidence-based system for their own practice varied, due in part to different learning experiences related to key aspects of the design.

Skilled facilitators such as Ms. Fina, could also ‘use’ this to plan their classroom observations and provide feedback to teachers on students’ progress throughout the year. In performing her role, Ms. Fina developed a trusting relationship with colleagues as the ostensive features of this design component intended, but at the same time, she regularly critiqued their work based on her deepening knowledge of SFA practices: What is more, the school’s principal used this facilitator’s knowledge of both the SFA instructional ideas, and the way such ideas were put into practice by teachers, to evaluate classroom instruction. Thus putting the design for evidence-based leadership into practice in this school was a balancing act, complicated by traditional egalitarian norms among teachers, and rules about teacher evaluations. Below Ms. Fina describes how over time, she learned to engage in this work and began to influence instructional change:

I make a note, like, if I observed. . . .you do Treasure Hunt. . . . To see if you’re implementing some of my ideas. There are some people, I can talk

myself blue in the face and I'll go in the next day and they're not doing anything the way I said it. Then I go in, I say, "Remember you've got to do this' and they say, "Um-hum, um-hum, um-hum.. And the next day I go in and they're still not doing it. Things like that become administrative. Because there's only so much authority I truly have. So I kind of just say, [to my principal,] "This one's up to you now. There's nothing else I can do."

I: How about with observations. It sounds like you provide teachers a lot of feedback.

Ms. Fina: I think that I've got the relationship that I can do it. There were teachers with rooms I would not go into three years ago, and they're not here anymore. I had a teacher, I'd go into her room, and she'd stand up and walk out in the hall. So I just stopped going into her room. She just didn't want to do the program. For the most part, I think they want the feedback.[S2002]

School norms and rules governing teachers' evaluation created challenges for the facilitator, even in Nightengale elementary our strong case of SFA implementation. In part because going beyond her allocated role of 'teaching' and supporting teachers' instructional change, to in essence evaluate teachers and assist in pressuring them to change, was difficult even for Ms. Fina, a strong facilitator.

This challenge was exacerbated by the SFA principals' generally weak knowledge of the design for instruction. Going into the school's fifth year of implementation, for instance, Ms. Fina's principal told us he did not understand the details of SFA well

enough to manage the evaluations on his own. Instead he used a ‘cheat sheet’ to evaluate his high poverty school’s reading program. We found that in Nightingale, this arrangement worked reasonably well, in part because the principal was serious about his willingness to delegate authority and the facilitator was very committed to the design. In other schools, where the Facilitator was less dynamic or the principal unable to provide this minimal level of support, the role division did not work as well. In Doris elementary school, the SFA Facilitator we interviewed told us flatly that there were classrooms she refused step inside, and teachers with whom she rarely spoke. As a result, there were clear limits to the leadership team’s ability to monitor the implementation of the design and advocate for low achieving students.

**Principals and Coaches in America’s Choice** . Leaders in America’s Choice schools collected many more forms of qualitative evidence than leaders in SFA. While only 4 of the SFA leaders we interviewed said that they examined samples of student’s written work, 24 of the AC leaders we interviewed in 2001 or 2002 said they examined this form of data. Twenty-one of these leaders provided strong and detailed accounts. As described earlier, in accordance with the AC design, leaders in our AC schools gathered, made sense of and used qualitative evidence based on a variety of different routines organized around the New Standards Performance Standards and a schedule in the leadership’s Monitoring Calendar. Information from source books, examples of writing in multiple genre and other forms of “data” were collected, evaluated publicly and used to both plan and continually revise/improve instruction.

Ms. Winn, the Design Coach at Redmond elementary explained the ‘what’ or content focus of the information collected in her school, as well as ‘how’ it was routinely used:

That’s [student assessment] also a big part of it. We look at student work. We look at class profiles for different genres. We look at commentary. We look at discipline logs and class meeting logs. With the assessment part, we use that to help our teachers plan for instruction for the next week, for the next three weeks, or whatever small group instruction they need for particular students. We can’t get in every classroom every day, but at least that’ll give us an idea. (Ms. Winn, Spring '02)

The AC design thus focused leaders’ and teachers’ attention on a wide range of information that was centered in classroom practice. Leaders would evaluate student work samples in their leadership team meetings and, in the first years of implementation, with teachers in their grade level meetings.

We actually get a monitoring calendar every month, telling them [teachers] what they need to bring in. We started genre studies at the end of April, so the month of May they’re going to be bringing in genre study samples. Or at the beginning of the year, they were working on narrative. We wanted them to bring in narrative. Whatever genre we were working on. . . . We ask them to bring in. . . a range [of work samples], so it’s not always the high students. . . because everyone needs help in different ways. . . . Sometimes we’ll have [teachers] actually do a sticky note [describing] what the strengths were for that work, what were the

weaknesses. Sometimes, we'll just have them share and talk about it and we'll look at the standard and see what elements and standards do the students meet (S2002).

The AC routines guided leaders and teachers in gathering specific kinds of information about the outcomes of instruction and created opportunities for educators to make sense of that information with colleagues (Daft & Weick, 1984; Fink & Resnick, 2001; Paredes Scribner, 1999). Not only did these routines produce an array of information for solving instructional problems, but leaders and teachers could judge the quality of learning in the school, professionally, based on commonly held external criteria specified in AC standards.

AC leaders also routinely gathered information about the quality of instructional performances from publicly displayed work and classroom observations. They used that information as one way to assess the quality of AC instruction and track the progress of the implementation of the AC design. Ms. Calm, one of the Coaches at Bonds elementary, describes the depth of her knowledge about AC instructional problems and how the leadership team put the AC system into practice through monthly collections of student work:

I: Do you think that teachers take what they learn in those meetings or in that modeling and actually put it into practice in instruction? Do you think they actually are able to use what they learn?

Ms. Calm: They all do, but some do it better than others.

I: How do you have a sense of what they're actually using?

Ms. Calm: The proof is in the student work. . . . For example, everyone has to do Response to Literature. They read a book and they have to respond to it. . . . The most difficult [standard task] is for our students to be able to have a book read or read a book, come up with a judgment, take that same judgment, and relate it to a life experience, and refer back to the book to verify that that life experience does indeed relate to the book. That's a very difficult skill. . . . Students have to submit Response to Literature to go on their Standards-Based Bulletin Board that's in the main hallway. And from those papers we can tell whether they're getting it or not, and what students seem to do a better job than others. And if it's not being taught to them, it's very evident in their papers.

(S2002)

This routine, school-wide assessment (Supovitz & Klein, 2003) allowed the team to discover which children were progressing, which students were falling behind, and which teachers had learned how to teach a particular skill. Importantly, this and other AC routines allowed teachers and leaders to use student work to evaluate the *depth* of change in classrooms, across the schools (Coburn, 2003; Cohen & Ball, 2007).

Leaders likewise evaluated the depth and breadth of the schools' improvement progress based on student-teacher performances in the action of classrooms. The literacy coach at Redmond Elementary described how she and other school or district professionals started focusing on the way students answer



questions during one portion of Writer’s Workshop (see also Barnes, Massell & Vanover (Under Review) for more on this case).

I: So are you where you think you should be with regard to the model?

LC Hart: In some areas. . . but there are some things that we need to work on. . . Author’s Chair for Writer’s Workshop. The kids’ responses need to really not just, ‘I like your story,’ but, ‘Why did you like the story? What part of the story did you like?’ We need to go deeper with the responses.

I: When you say “deeper,” what do you mean? What kinds of things do teachers have to do differently?

LC Hart: They have to model it, because the kids don’t know how to respond. That’s something that we’re seeing across all the schools; that the kids are responding, but they’re not responding. . . .because they’re saying things they know the teacher wants to hear. And now we want them to make it more real for them to internalize those types of things. [S2002]

In this instance, implementation of the AC design focused leaders’ attention directly on the ways that students responded to the moment-to-moment operation of the reform program. Sixteen of the AC leaders we interviewed told us they would observe what individual students were learning when they walked into classrooms, as opposed to 4 leaders in SFA. This pattern of leaders focusing on a range of student performances held across our AC sites. Leaders in AC gathered evidence about students and instruction in an up close’ manner through a process that was

embedded in both classroom practice and external, commonly held standards of quality.

The AC design intended that quantitative data from the New Standards Reference Exam be streamed into this system. However, while the test was used in 16 of the schools we studied, only 4 of the school leaders we interviewed provided detailed descriptions of using these data. By the 2002 spring interviews, only one school planned to use the Reference Exam in the coming school year. Instead, schools were either using, or planning to use district assessments, in part due to the cost, and in part because leaders were overwhelmed by the amount of testing. As a result, this element of the design was replaced with data from district and state assessments that were not always aligned with the AC's performance standards. Several AC coaches and several teachers complained about nature of the data from these latter tests: the information they generated was more “distant” from practice than the data the Reference Exam generated. None of the AC schools we studied in 2002 implemented the type of school-wide quantitative assessments that were in evidence in all of our SFA case study schools. As a result, they were not able to focus on student reading skills in the same way there were able focus on improvements in writing.

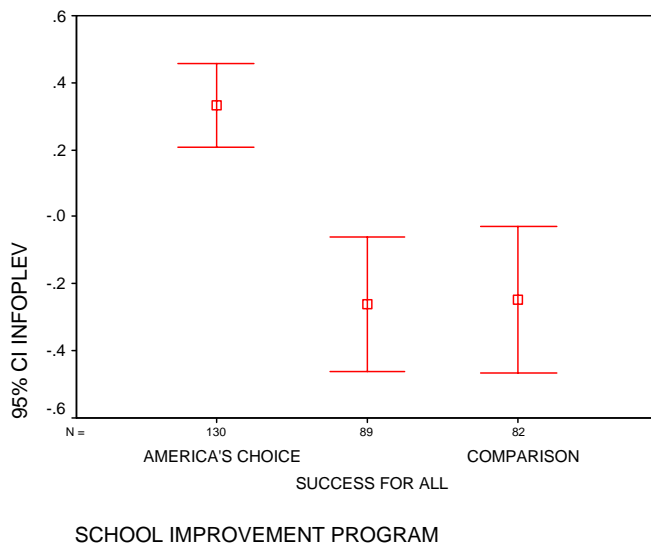
### **Quantitative Descriptions of Evidence-Based Leadership Practice**

We used three quantitative measures to describe the way school leaders in our schools gathered, worked with and used various sources of evidence about instruction in their schools. These factors make up some of the key elements and processes in the two designs' “systems of evidence-based leadership” that enactors in our schools were more or less able to put into action—i.e. to activate in their schools. The descriptive

quantitative evidence that follows, while not a one to one correspondence with our qualitative data, complements it, and is suggestive of the trends we found.

First, based on our qualitative data, we assumed that AC school leaders would report using a wide range of evidence in planning or evaluating their schools' progress. Our quantitative descriptions support this conjecture. AC school leaders in our sample reported using information from planning and evaluation from a variety of sources to a greater extent than leaders in SFA and comparison schools: for example they discussed how they focused on results from standardized assessments and qualitative assessments such as student conferences, portfolios, samples of student work, and classroom observations. The difference in means between AC school leaders and the leaders in SFA and comparison schools was significant on this INFO-USE measure ( $p < .001$ ). But the means between SFA leaders and the comparison leaders were not significantly different (Table 2.3 in Appendix A).

**Figure 3. INFO-USE by Program**



To better understand this finding we examined item-level differences between the programs (Table 3 below and Figure 1.1. in Appendix A). The INFO-USE scale consists of five items, which represent five different types of information. Over half of AC leaders reported using all five types of information to a great extent. Unlike this consistent pattern among AC leaders, SFA leaders' reports were somewhat inconsistent. About 67% of SFA leaders reported using standardized, curriculum-referenced test results 'to a great extent' while the percentages for using the other four types of information to a great extent ranged between 30-40% among SFA leaders. (See Table 3., below and Figure 1.1., Appendix A.)

Table 3. School leaders' ratings of the extent to which leaders using information for planning and evaluation

| Types of information                             |      | To what extent do you use |                   |                      |                   |
|--|------|---------------------------|-------------------|----------------------|-------------------|
|  |      | Not at all                | To a small extent | To a moderate extent | To a great extent |
| Standardized, curriculum-referenced test results | AC   | .8                        | 4.2               | 31.9                 | 63.0              |
|  | SFA  | 2.5                       | 6.2               | 23.5                 | 67.9              |
|  | Comp | 5.5                       | 6.8               | 38.4                 | 49.3              |
| Informal assessments by teachers**               | AC   | 1.7                       | 5.9               | 31.4                 | 61.0              |
|  | SFA  | 2.5                       | 17.3              | 39.5                 | 40.7              |
|  | Comp | 5.6                       | 13.9              | 43.1                 | 37.5              |
| Learning or curriculum standards**               | AC   | 0                         | 5.1               | 28.0                 | 66.9              |
|  | SFA  | 2.5                       | 14.8              | 45.7                 | 37.0              |
|  | Comp | 1.4                       | 11.1              | 38.9                 | 48.6              |
| Samples of student academic work***              | AC   | .8                        | 2.5               | 34.7                 | 62.0              |
|  | SFA  | 3.8                       | 27.5              | 38.8                 | 30.0              |
|  | Comp | 5.5                       | 20.5              | 42.5                 | 31.5              |
| Your own or                                      | AC   | .8                        | 7.4               | 38.0                 | 53.7              |

|                                |      |     |      |      |      |
|--------------------------------|------|-----|------|------|------|
| others' classroom observation* | SFA  | 2.5 | 18.5 | 42.0 | 37.0 |
|                                | Comp | 4.2 | 16.7 | 43.1 | 36.1 |

- The numbers represent the percentages of school leaders' ratings within each program

- \* p<.05; \*\* p<.01; p<.001 for Chi-Square Test

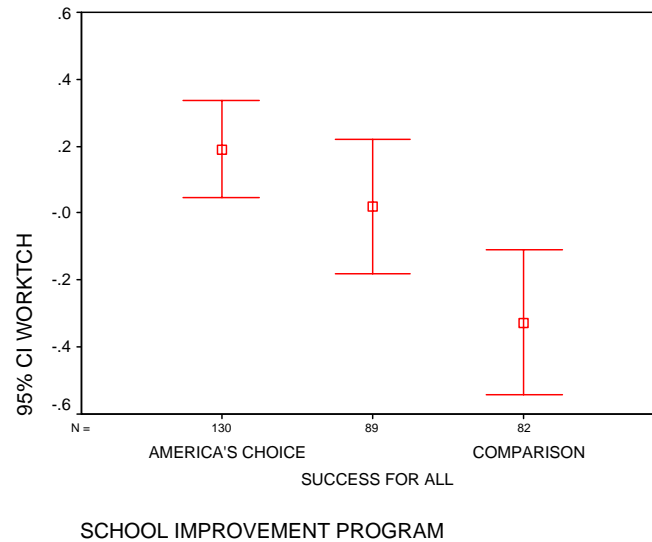
- Disproportionally more AC leaders report using various types of information to a great extent. An exception is the use of standardized, curriculum-referenced test results. SFA leaders' rating on this item is as high as AC leaders'.

Thus, the quantitative findings tend to complement our qualitative findings. Teams of SFA leaders across schools focused on the designs' quarterly assessment data. AC leaders, however, focused on a wider variety of evidence—especially performance standards, student work, teachers' formative assessments and classroom observations.

Based on our qualitative data and descriptions of leaders practice we also assumed that though the emphasis might vary, leaders in both the SFA and the AC designs would observe teachers and use student work or standardized test results to advise them on instructional practices more than leaders in schools not using one of the designs. We assumed that one of the ways that the designs added value was to change the systems of routines that organized leaders' work and by doing so, de-privatize instructional practice (Camburn et al., 2003). The quantitative findings show that compared to the leaders in comparison schools, on average, AC school leaders and SFA leaders reported a higher level of working directly with teachers in these ways (Figure 4). But the mean difference between the two designs is not statistically significant (Table 2.3. in Appendix A).<sup>8</sup>

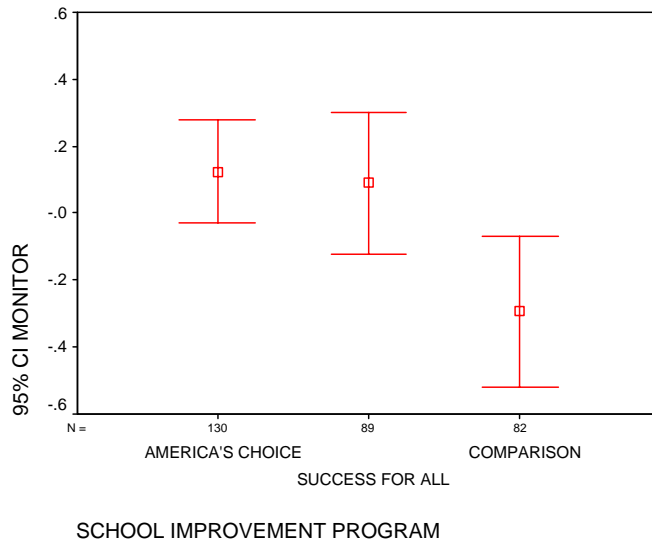
#### Figure 4. WORKTCH by Program

<sup>8</sup> See the post-hoc test in our ANOVA analysis (see [Table 2.3 in Appendix A](#)).



On average, AC and SFA school leaders in our sample also reported engaging in various monitoring activities to a greater extent than their counterparts in comparison schools; that is, engaging in practices that monitor the improvement efforts in the schools (Figure 5). Compared to the leaders in comparison schools, AC and SFA leaders reported that they more often monitor curriculum and instructional practice, observe in classrooms, and evaluate teachers with criteria directly related to improvement efforts in their schools. Again, the means for this scale were not significantly different between AC and SFA school leaders (Table 2.3. in Appendix A).

Figure 5. MONITOR by Program



While these descriptions show that leaders in both designs use elements of what we have labeled “evidence-based” practice, they do not show as our qualitative findings do, the differences in the nature of those practices. As discussed above, SFA leaders tended to work with teachers for short time periods. They moved quickly from teacher to teacher to monitor pacing and content coverage. AC leaders enacted a different strategy. Their efforts tended to focus on evaluating student work according to the standards and working intensively with teachers for longer periods of time on new instructional practices.

## CONCLUSION

Our study shows that the CSR designs were not mechanisms that instantaneously transformed local school environments. Their resources could not renovate leaders and teachers’ practice unless they were used well in school professionals’ daily work (Cohen & Ball, 1999; Cohen et al., 2003; Correnti & Rowan, 2007). School improvement was a learning process based on more than general, reform principles; enactors were given concrete and specific guides to action. But fidelity was not constructed through mindless

adherence to standard operating procedures. Instead, skilled use of the designs required that participants understand what they were doing and be capable of using design resources to evaluate their efforts (Barnes, Khorsheed, De Los Rios, and Correnti, 2006). In the schools we studied, performing the routines required by the designs' ostensive elements in the context of schools posed many challenges. Leaders had to learn the changes to teaching and leading called for by the designs; they had to learn new ways of working with other school professionals to put these changes into practice; and they had to learn how to use new forms of assessment to monitor the improvement effort and improve the quality of professionals' work.

Thus, using an interpretive conceptual frame and triangulating our data we have argued that evidence-based leadership is not solely an intellectual exercise. It is a learned system of practice that, ideally, builds resources iteratively as participants learn what to do and how to do it well. We found that much of the variation in implementation of such systems, within and across designs, depended in part on three interrelated resources: leaders' knowledge of instruction, their motivation to engage in school improvement, and the nature of the tools, guided routines or learning opportunities structured into the designs. Importantly, the school context and conditions of work in which leaders were using the designs for improving their practices were also critical factors in shaping the quality of implementation. While the designs provided new tools and guides to practice they also made new demands on school leaders.

Our findings show that the systems of leadership that emerged across the AC and SFA schools in our study did vary by design providing strong evidence that change in leadership practice had occurred. We found that many ostensive features of CSR



programs did indeed shape leaders' performances in ways that were roughly consistent with the designs' intentions. The change process for school leaders across both designs was facilitated by formal and informal opportunities to learn and to come to understand the instructional core of the design. Leaders reported learning off-site as well as in their schools and from enacting the designed leadership practices. Thus, some leaders were able to acquire new knowledge and skills over time from their practice as it was shaped by the design.

In general, principals' reports of change tended to be weaker than the extensive changes and deep understanding of the designs described by facilitators and coaches across both designs. But principals' reports varied by design in AC and SFA schools as well. AC principals reported considerably more change in their practice that, in their account, resulted in a greater focus on classroom instruction than did SFA principals. SFA principals, at the time of our study, reported far fewer opportunities to learn how to lead instruction and manage the demands of evidence-based leadership practice from the daily operation of the design.

The distribution of leadership authority, tasks and knowledge (Spillane, 1999) by the designs' ostensive features and by how leaders implemented them in the school context, influenced the extent to which elements of the respective designs could work together in a mutually reinforcing system. Some principals felt they could not, or were not motivated to set aside their myriad other pressing responsibilities to learn about the designs instructional core. CSR 'teacher leaders' worked in the context of traditional norms of teacher autonomy and union rules about authority over assessment or evaluation of teaching. Thus while the designs' intent and positional authority warranted principals'

monitoring of the designs' instruction, some of them did not have the knowledge to do this task well. On the other hand, 'teacher leaders' serving as CSR coaches did not have positional or normative authority for the task, but they were well situated to engage in knowledgeable critique of instruction compared to the designs' precepts. Some skilled coaches tended to gain the respect and authority of teachers while keeping the principal informed. But there were schools where the leadership team was so weak that elements of the design's ostensive system were never pulled together into a coherent structure. The extent to which school leadership teams could put these elements of the design together into a coordinated system influenced the quality of the designs' implementation.

At a time when, as Raudenbush (2005) emphasizes, all roads lead to instructional improvement, our paper shows the importance of school level leadership and evidence-based practice for supporting these efforts. As other researchers show, (e. g. Coburn, 2003; Elmore, 2000; Fink & Resnick, 2001; Rowan & Miller, 2007; Spillane, Halverson, & Diamond, 2004) changing classroom instruction requires changing leaders' practice. Our findings also offer guidance for designing these systems for future change programs. They highlight the importance of school-wide systems of qualitative and quantitative data and emphasize the benefits of using these assessments in iterative cycles of improvement to support implementation across a long period of time. The findings we have presented also illustrate the importance of professional development for leaders as a key part of this improvement effort. When the designs' resources were constructed well, professionals learned how to build on the routines such design imported into their school and use the tools the design offered to guide improvement.

## **References**

- Allensworth, E., Correra, M., & Ponisciak, S. (2008). *From high school to the future: ACT Preparation--too much, too late*. Chicago: Consortium on Chicago School Research at the University of Chicago.
- Barnard, C. I. (1938). *The functions of the executive*. Cambridge, MA: Harvard University Press.
- Becker, M. C. (2004). Organizational routines: A review of the literature. *Industrial and Corporate Change*, 13, 643-678.
- Berends, M. (2004). In the wake of A Nation at Risk: New American Schools' private sector school reform initiative. *Peabody Journal of Education*, 79(1), 130-163.
- Berends, M., Bodilly, S. J., & Kirby, S. N. (2002). *Facing the challenges of whole-school reform: New American Schools after a decade*. Santa Monica, CA: Rand: Education.
- Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas Accountability System. *American Educational Research Journal*, 42(2), 231-268.
- Braybrook, D. (1964). The mystery of administrative success re-examined. *Administrative Science Quarterly*, 1964(8), 533-560.
- Camburn, E., Rowan, B., & Taylor, J. (2003). Distributed leadership in schools: The case of elementary schools adopting comprehensive school reform models. *Educational Evaluation and Policy Analysis*, 25(4), 347-373.
- Camburn, E., Rowan, B., & Taylor, J. (2004). Distributed leadership in schools: The case of elementary schools adopting comprehensive school reform models. *Education Evaluation and Policy Analysis*, 25(4), 347-375.
- Coburn, C. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, 32(6), 3-12.
- Coburn, C., & Talbert, J. E. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education*, 112(August 2006).
- Cohen, D. K., & Ball, D. L. (1999). *Instruction, capacity and improvement* (CPRE Research Report Series RR-43). Philadelphia, PA: Consortium for Policy Research in Education: University of Pennsylvania.
- Cohen, D. K., Raudenbush, S. W., & Ball, D. L. (2003). Resources, instruction and research. *Educational Evaluation and Policy Analysis*, 2, 119-142.
- Correnti, R., & Rowan, B. (2007). Opening up the black box: Literacy instruction in schools participating in three comprehensive school reform programs. *American Educational Research Journal*, 44(2), 298-338.
- Crandall, D. P., & Loucks, S. F. (1983). *A roadmap for school improvement: Executive summary of the study dissemination efforts supporting school improvement. People, policies, and practices: Examining the chain of school improvement* (Vol. 10). Andover, MA: The Network.

- Daft, R. L., & Weick, K. E. (1984). Toward a model of organizations as interpretation systems. *Academy of Management Review*, 9(2), 284-293.
- Datnow, A., & Stringfield, S. (2000). Working together for reliable school reform. *Journal of Students Placed at Risk*, 5(1-2), 183-204.
- Desimone, L. (2002). How can comprehensive school reform models be successfully implemented? *Review of Educational Research*, 72(3), 433-450.
- Diamond, J. B., & Spillane, J. (2004). High-stakes accountability in urban elementary schools: Challenging or reproducing inequality. *Teachers College Record*, 106(6), 1145-1176.
- Dutton, J., Worline, M., Frost, P., & Lilius, J. (2006). Explaining compassion organizing. *Administrative Science Quarterly*, 51(2006), 59-96.
- Elmore, R. F. (2000). *Building a new structure for school leadership*. Washington, DC: Albert Shanker Institute.
- Elmore, R. F., Peterson, P. L., & McCarthy, S. J. (1996). *Restructuring in the classroom: Teaching, learning and school organization*. San Francisco: Jossey-Bass.
- Feldman, M. S. (2003). A performative perspective on stability and change in organizational routines. *Industrial and Corporate Change*, 12(4), 727-752.
- Feldman, M. S., & March, J. G. (1981). Information in organizations as signal and symbol. *Administrative Science Quarterly*, 26(2), 171-186.
- Feldman, M. S., & Pentland, B. T. (2003). Reconceptualizing organizational routines as a source of flexibility and change. *Administrative Science Quarterly*, 48(1), 94-118.
- Feldman, M. S., & Rafaeli, A. (2002). Organizational routines as sources of connections and understandings. *Journal of Management Studies*, 39, 309-331.
- Fink, E., & Resnick, L. B. (2001). Developing principals as instructional leaders. *Phi Delta Kappan*, 82(8).
- Firestone, W. A., & Corbett, H. D. (1988). Planned organizational change. In N. Boyand (Ed.), *Handbook of research on educational administration*. White Plains, New York: Longman.
- Hutchins, E. (1995a). *Cognition in the wild*. Cambridge, MA: The MIT Press.
- Hutchins, E. (1995b). How a cockpit remembers its speeds. *Cognitive Science*, 19, 265-288.
- Leavitt, B., & March, J. G. (1988). Organizational learning. *Annual Review of Sociology*, 14, 319-340.
- March, J. G., & Simon, H. A. (1958). *Organizations*. New York: Wiley.
- Murphy, J., Elliott, S. N., Goldring, E., & Porter, A. (2007). Leadership for learning: A research-based model and taxonomy of behaviors. *School Leadership & Management*, 27(2), 179-201.
- National Research Council (Ed.). (2003). *Assessment in support of instruction and learning: Bridging the gap between large-scale and classroom assessment: Workshop report*. Washington, D. C.: The National Academies Press.

- O'Day, J. A., & Smith, M. S. (1993). Systemic reform and educational opportunity. In S. H. Fuhrman (Ed.), *Designing coherent educational policy: improving the system*. San Francisco: Jossey-Bass.
- Paredes Scribner, A. (1999). Using student advocacy assessment practices. In P. Reyes & J. D. Scribner & A. Paredes Scribner (Eds.), *Lessons from High-Performing Schools: Creating Learning Communities* (pp. 169-187). New York: Teacher's College Press.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, D. C.: National Academy Press.
- Peurach, D. (2005). *Designing and managing comprehensive school reform: The case of Success For All.*, University of Michigan, Ann Arbor.
- Raudenbush, S. W. (2005). Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher*, June/July, 25-31.
- Resnick, L. B. (1995). From aptitude to effort: A new foundation for our schools. *Daedalus*, 124, 55-62.
- Resnick, L. B., & Glennan, T. K. (2002). Leadership for learning: A theory of action for urban school districts. In A. T. Hightower & M. S. Knapp & J. A. Marsh & M. W. McLaughlin (Eds.), *School districts and instructional renewal*. New York: Teachers College Press.
- Rowan, B., Camburn, E., & Barnes, C. (2004). Benefiting from Comprehensive School Reform: A review of research on CSR implementation. In C. T. Cross (Ed.), *Putting the pieces together: Lessons from comprehensive school reform research* (pp. 1-52). Washington, DC: George Washington University Press.
- Rowan, B., Camburn, E., & Correnti, R. (2004). Using teacher logs to measure the enacted curriculum in large-scale surveys: Insights from the Study of Instructional Improvement. *Elementary School Journal*, 105, 75-102.
- Rowan, B., & Correnti, R. (2007). Opening up the black box: Literacy instruction in schools participating in three comprehensive school reform programs. *American Educational Research Journal*, 44, 298-338.
- Rowan, B., & Miller, R. J. (2007). Organizational strategies for promoting instructional change: Implementation dynamics in schools working with comprehensive school reform providers. *American Educational Research Journal*, 44(2), 252-297.
- Scribner, J. D. (1999). Creating learning communities for high-performing Hispanic students: A conceptual framework. In P. Reyes & J. D. Scribner & A. Paredes Scribner (Eds.), *Lessons from High-Performing Hispanic Schools: Creating Learning Communities* (pp. 188-210). New York: Teachers' College Press.

- Shepard, L. A. (2000). The role of classroom assessment in teaching and learning. In V. Richardsen (Ed.), *Handbook of Research on teaching* (4th ed.). Washington, D. C.: American Educational Research Association.
- Skrla, L., & Scheurich, J. J. (2001). Displacing deficit thinking in school district leadership. *Education and Urban Society*, 33(3), 235-259.
- Skrla, L., Scheurich, J. J., Garcia, J., & Nolly, G. (2004). Equity audits: A practical leadership tool for developing equitable and excellent schools. *Education Administration Quarterly*, 40(1), 133-161.
- Spillane, J., Diamond, J. B., & Jita, L. (2004). Leading instruction: The distribution of leadership for instruction. *Journal of Curriculum Studies*.
- Spillane, J., Halverson, R., & Diamond, J. (1999). *Distributed leadership: Toward a theory of school leadership practice*: Northwestern University.
- Spillane, J., Halverson, R., & Diamond, J. (2004). Towards a theory of leadership practice: A distributed perspective. *Journal of Curriculum Studies*, 36(1), 3-34.
- Spillane, J., & Sherer, J. Z. (2004, April, 2004). *A distributed perspective on school leadership: leadership as "stretched over" people and place*. Paper presented at the American Educational Research Association, San Diego.
- Stein, M. K., & Nelson, B. (2003). Leadership content knowledge. *Educational Evaluation and Policy Analysis*, 25(4), 423-448.
- Stringfield, S. (1995). Attempting to enhance students' learning through innovative programs: The case for schools evolving into high-reliability organizations. *School Effectiveness and School Improvement*, 6(1), 67-96.
- Supovich, J., & Klein, V. (2003). *Mapping a course for improved student learning: How innovative schools systematically use student performance data to guide improvement*. Philadelphia: Center for Policy and Research in Education.
- Supovich, J., & Poglinco, S. M. (2001). *Instructional leadership in a standards-based reform*. Philadelphia, PA: Consortium for Policy Research in Education: University of Pennsylvania.
- Supovitz, J., & Klein, V. (2003). *Mapping a course for improved student learning: How innovative schools systematically use student performance data to guide improvement*. Philadelphia: Center for Policy and Research in Education.
- Walsh, J. P. (1995). Managerial and organizational cognition: Notes from a trip down memory lane. *Organizational Science*, 6(3), 280-321.
- Walsh, J. P., & Rivera Ungson, G. (1991). Organizational memory. *The Academy of Management Review*, 16(1), 57-91.
- Weick, K. E., & Roberts, K. (1993). Collective mind in organizations--heedful interrelating on flight decks. *Administrative Science Quarterly*, 38(3), 357-381.