WCER Working Paper No. 2009-2
March 2009

District-Wide Reform of Mathematics and Science Instruction: Case Studies of Four SCALE Partner Districts

William H. Clune
University of Wisconsin Law School/
Wisconsin Center for Education Research
University of Wisconsin–Madison
whclune@wisc.edu

WCER
Wisconsin Center for Education Research
School of Education • University of Wisconsin–Madison • http://www.wcer.wisc.edu/

WCER working papers are available on the Internet at http://www.wcer.wisc.edu/publications/workingPapers/index.php. Recommended citation:

> Clune, W. H. (2009). *District-wide reform of mathematics and science instruction: Case Studies of four SCALE partner districts* (WCER Working Paper No. 2009-2). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved [e.g., March 30, 2009,] from http://www.wcer.wisc.edu/publications/workingPapers/papers.php

The research reported in this paper was supported by a grant from the National Science Foundation to the University of Wisconsin–Madison (EHR 0227016) for a Mathematics and Science Partnership project called System-wide Change for All Learners and Educators (SCALE), and by the Wisconsin Center for Education Research, School of Education, University of Wisconsin–Madison. Any opinions, findings, or conclusions expressed in this paper are those of the author and do not necessarily reflect the views of the funding agencies, WCER, or cooperating institutions.

District-Wide Reform of Mathematics and Science Instruction: Case Studies of Four SCALE Partner Districts[1]

William H. Clune

This paper is a synthesis of case studies of four districts that implemented multifaceted reforms aimed at offering rigorous instruction in mathematics and science for all students. The districts—urban or urbanizing and ranging in size from medium to very large—were partners with three universities in System-wide Change for All Learners and Educators (SCALE), a National Science Foundation (NSF)-funded Math and Science Partnership (MSP) project. Rigorous instruction for all students requires depth and breadth of impact. To achieve these ambitious goals, the districts adopted and adapted a combination of policies (formal and informal) under a common theory of action jointly articulated and developed by and for the partnership.

This paper describes the partnership, the common theory of action, evaluative criteria derived from the theory, implementation in each district, and the effectiveness of the reforms measured against the theory of action. The paper concludes by suggesting models for effective district policy in other districts.

The SCALE Partnership

Funded in 2002, SCALE was a comprehensive MSP project that included four major urban school districts—the Denver Public Schools (DPS), Los Angeles Unified School District (LAUSD), Madison Metropolitan School District (MMSD), and Providence (RI) Public School District (PPSD)—and four universities—the University of Wisconsin–Madison (UW-Madison); California State University, Dominguez Hills (CSUDH); California State University, Northridge (CSUN), and the Institute for Learning (IFL) of the University of Pittsburgh, which was a partner during the period when much of the data for the case studies was collected.[2] LAUSD, DPS, and PPSD were affiliated with IFL prior to SCALE. MMSD was introduced to IFL through SCALE, and UW-Madison played a strong coordinating role for MMSD similar to that played by IFL for the other three districts.

Under the partnership, each district made a commitment to a common vision and theory of action for instructional reform. Agents of the partnership external to the districts (universities, other districts) participated in district reform efforts through a process that came to be thought of

---

[1] Special thanks to SCALE colleagues Susan Millar, Eric Osthoff, and Lauren Resnick for detailed comments on this paper (not necessarily implying agreement with its analysis and conclusions).

[2] The Institute for Learning was founded in 1995 as a partnership of school districts committed to standards-based education and system-wide reform. It serves as a liaison between its parent institution, the Learning Research and Development Center at the University of Pittsburgh, and educators in school systems nationwide. It bridges the domains of research and practice by conveying to educators the best of current knowledge and research about learning processes and principles of instruction. IFL believes that for education reform to be broadly effective and equitably distributed, the unit of change must be the district. As a result, the programs IFL develops are designed not for schools or individuals but rather for entire districts. Disagreements over the direction of SCALE eventually led to the withdrawal of IFL from the partnership. Those issues are beyond the scope of this paper (with its focus on districts) but are part of the Building a Partnership branch of SCALE research and evaluation.

as *co-construction*—that is, a broad set of collaborative activities (planning, training, product design, piloting, and evaluation) with a wide range of district actors (e.g., top management, departments of curriculum and instruction, principals, coaches, teachers) (Clune, 2005).

The SCALE Theory of Action

Weiss (1998) posited that the theory of change of a program (here referred to as *theory of action*) has two components—implementation theory (i.e., the intended implementation) and program theory (i.e., the anticipated behavioral responses of those affected by the implementation)—both hypothesized to lead to the achievement of the program's intended outcomes. The SCALE theory of action had all of the elements specified by Weiss—intended outcomes, a design for implementation, and anticipated behavioral responses.

With respect to intended outcomes, SCALE had a clear, overarching goal of high-quality instruction for all students. As stated by IFL Director Lauren Resnick (2006) in a draft design paper for SCALE:

> The problem that SCALE is addressing is how to take ambitious ideas for mathematics and science teaching reform "to scale"—across all grade levels and all groups of students in our urban districts. We are not trying to create a few schools that are models of successful STEM [science, technology, engineering, and mathematics] teaching and learning, but rather are aiming to make effective high-demand teaching and learning the rule, rather than an applauded exception, in each SCALE district. It is difficult to overestimate the pedagogical and organizational challenges embedded in this goal. Helping teachers within a school to implement effective instructional practice is a major challenge, and the idea of doing so across a significant number of schools, with the diverse and changing student and teacher populations found in major urban districts, is daunting. (p. 2)

Districts pursued this goal through two major reform efforts, rooted in the following two SCALE goals:[3]

1. *Goal 1: Improve core instruction.* Improve all regular math and science courses through district policies governing curriculum, professional development, monitoring of instructional quality, and accountability, all aligned with each other and with the overarching goal.

2. *Goal 2: Science immersion for all students.* Develop and implement immersion STEM learning experiences to ensure that every student experiences the process of engagement in an extended (e.g., 4-week) scientific investigation at least once a year.

The following discussion describes the implementation design for Goal 1 as developed early in the partnership. The design for Goal 2 emerged later and contained parallel elements, as described subsequently.

[3] Goal 3, which dealt with the contributions of faculty in institutions of higher education, facilitated both Goal 1 and Goal 2.

Professional development of teachers. For professional development, SCALE adopted an IFL model that relied on training of coaches and lead teachers who in turn would train teachers. Following is a description of the model in the SCALE Year 3 implementation plan:

> IFL/immersion facilitators provide special training in effective facilitation; coaches and lead teachers will often be coached "at elbow" as they begin to do the work. "At elbow" or shadow coaching is an essential aspect of our strategy for taking SCALE to scale. It bypasses the typical "cascade" effect,[4] and brings the learning directly to the classroom. The philosophy of creating an environment in which instructional coaches and lead teachers collaborate is a cornerstone of the theory of action in all four SCALE districts. All four districts have outlined their plans for recruiting and training coaches and lead teachers.

Curriculum. The SCALE theory of action specified curriculum guidance that conformed to state standards and encouraged academic rigor. In the words of a SCALE planning document submitted by IFL:

> Each SCALE district is aiming for instructional programs that accord with our principle of *Academic Rigor in the Thinking Curriculum*. This principle calls for teaching programs that include important conceptual content in accord with the state's standards *and at the same time* engage students in high levels of cognitive engagement with the content. We call this *Teaching and Learning on the Diagonal.* Whatever name we give it, commitment to the academic rigor principle calls on each district to examine the content of its teaching programs to make clear assessments of the extent to which the content being taught is rigorous, important, aligned and demands high levels of cognitive functioning.

One model for curriculum guidance was the designed curriculum (or program) defined by Resnick as follows in the draft design paper:

> By a "designed" program, we mean one that has been built to careful specifications, usually theory-based, and tested in use. Designed programs are usually crafted to fit into the time actually available for instruction. They would typically invite considerably less "picking and choosing" by teachers than traditional textbooks. Designed programs may use traditional written materials (textbooks, workbooks, etc.) or be computer-based.

A mandated curriculum simplifies policy alignment. Professional development can be organized around a single curriculum rather than trying to serve multiple textbooks and supplements. Curriculum depth can be enhanced through the problems, learning activities, and assessments embedded in instructional units. Three SCALE districts—DPS, MMSD, and PPSD—used NSF-supported reform curricula that qualified as designed (e.g., in mathematics, *Everyday Mathematics*, *Investigations*, and *Connected Mathematics*; and in science, the Full Option Science System [FOSS]); but more traditional curricula may qualify as designed in the

---

[4] *Cascade* refers to a train-the-trainer model through which an external agent trains district staff who are responsible for professional development sessions. The alternative discussed in the quotation is training of coaches and lead teachers who provide and facilitate site-based professional development in schools.

same sense (e.g., *Saxon Algebra*). In the absence of a single mandated curriculum, rigorous curriculum content was to be pursued through other forms of instructional guidance, such as instructional guides and student assessments that would influence choice of textbooks, supplementary materials, and the selection and emphasis of topics for teaching.

Monitoring. Monitoring of instructional quality (the *enacted curriculum*) was intended to provide feedback for possible modifications of other strategies, like professional development (SCALE, 2003):

> Monitoring systems can be quite formal (for example, periodic assessments of students with data going into a central data analysis system) or more informal (for example, periodic school visits that include classroom observations and analysis of student work products). In either case, their goal is formative and diagnostic at a system level. That is, they allow senior administrators (superintendents and chief academic officers; supervisors of principals) to gauge how schools are doing; they may also allow principals to gauge how teachers or groups of teachers are doing, in their efforts to teach all students. They examine achievement, and teaching performance . . . systematically and frequently, not with the intent of rewarding or punishing but rather of determining where further support or training is needed. (p. 16)

Various other methods of monitoring instructional quality were actively considered by the SCALE leadership and district staff (e.g., the Survey of Enacted Curriculum),[5] but as time progressed the major efforts gravitated toward quarterly assessments of student achievement. Student assessments provide information about the content and quality of instruction to the extent that student performance is or is not correlated with *opportunity to learn* (the enacted curriculum). For example, difficulties encountered by students on the assessments provide clues about which topics should be strengthened. Limitations of using interim assessments as a vehicle for monitoring and improving teaching are discussed in the section on PPSD later in this paper.

Accountability. SCALE lacked a fully developed theory of action for accountability, instead relying on the accountability mechanisms already in place in the districts. The intent was to use accountability as a source of leverage for the other elements of the theory of action aimed at instructional improvement. In practice, interim assessments were the most influential type of accountability, resulting in a merger of monitoring and accountability.

System management and organizational learning. Initially, the SCALE theory of action did not include a system dimension concerned with coordination, management, and organizational learning. The IFL theory of action did include this component (see below), and its importance within SCALE soon became evident. System management is the invisible hand that keeps the separate reform components and organizational actors aligned with ultimate goals. It is the mechanism of distributed leadership insuring buy-in and support from all levels of the organization (Spillane, Halverson, & Diamond, 2001). And it is the instrument of organizational learning at every level, providing information, feedback on consequences, and a process for reactive problem solving. Complex reforms like those in SCALE, aided by external partners and

[5] http://seconline.wceruw.org/secWebHome.htm

unfolding as design experiments (problem-oriented research and development), obviously require sophisticated guidance.

In the IFL theory of action, the system management and organizational learning component was built on a culture of continuous learning, two-way accountability, and nested learning communities:

> Commitment to continued learning on the part of all professionals in the system is a central feature of the IFL's district reform efforts. Schools need to become places of learning for teachers as well as for students, and principals have a primary responsibility for bringing this about. In these nested learning communities, principals should be accountable to teachers in the same way that teachers are accountable to students; that is, principals have a responsibility to establish effective learning opportunities for teachers. At the same time, just as students must take advantage of the learning opportunities that teachers provide if they are to achieve, so teachers must take advantage of the professional learning opportunities provided and expected in their schools. Principals and senior district administrators should be in a similar relationship of two-way accountability for continuous learning and school improvement. (Resnick & Glennan, 2002, p. 166)

According to Osthoff (2005), IFL staff typically focused on six areas to support and guide district actors in enacting their theory of action for improving teaching and learning (see also Resnick & Glennan, 2002):

1. Deep, visible engagement of the superintendent in instructional improvement;

2. Development of instructional leadership skills of first-line managers (i.e., principals and assistant principals);

3. Provision of tools to help principals and district staff develop their own and teachers' instructional capabilities;

4. Focus on core academic skills, typically literacy and mathematics;

5. On-the-job coaching of principals, district leaders, and teachers; and

6. Use of school, department, and leadership learning plans to foster nested learning communities and two-way accountability.

Elements of SCALE theory of action in Goal 2 science immersion units. A theory of action for Goal 2 science immersion units emerged over time, as best illustrated by the case study of science immersion in LAUSD summarized later in the paper. But Goal 2 was influenced by Goal 1 and had parallel elements that can be introduced and briefly described here.

Early on, the Goal 2 team adopted the IFL *principles of learning*[6] as an overarching goal and issued a paper merging the original immersion model with these principles (Schunn, Millar, & Lauffer, 2006). Immersion units clearly qualified and functioned as a designed curriculum, containing a detailed model of instruction as well as embedded problems and exercises. Alignment with district standards (mainly scope and sequence) became a major part of the immersion enterprise. Professional development was built around workshops providing inquiry-based training of facilitators and teachers. System coordination was achieved through an innovative version of co-construction and co-implementation among external partners and district personnel.

The Role of the Partnership and District Theories of Action

Although previously touched upon in this paper, partnership requires its own discussion. In the first piece I wrote on the subject (Clune, 2005), I portrayed the SCALE theory of action as follows:

> The partnership theory of action has an influence on →
> the district theory of action, embodied in →
> active district instructional guidance within the SCALE mission, which leads to →
> improvement of schools, instruction, and student learning.

Under this model, the "programs" that are typically the focus of evaluation under a theory-of-action approach "operate within and are managed by the districts. The question then becomes how to describe the theory of action through which the partnership influences these programs in a way that ultimately produces the desired outcomes" (Clune, 2005, p. 2). That influence was anticipated to occur through wide-ranging activities of co-construction, described in that first paper as follows:

> [C]o-construction activities . . . tend to fall into one of four categories: training, product design, joint planning (or review), and co-piloting. . . . [t]he activities are carried out at every level of the system, from top management to preparing and coaching teachers. (Clune, 2005, p. 5)

These ideas were formulated after months of intensive discussion and learning about how the partnership would work among SCALE implementers and researchers.

If anything, the 2005 paper may have given excessive weight to partnership influence. Evaluation of partnership influence—as documented by the case studies—is vital and was the focus of SCALE's Building a Partnership Team. Partnership in this sense refers to assistance provided by SCALE partners to the district. But the policies supported by the partnership were executed within the district, mostly through district effort and according to its local theory of action. In fact, the SCALE theory of action for district policy was developed as a distillation of

[6] According to the IFL (2007), the IFL principles of learning are "condensed theoretical statements summarizing decades of learning research" and entail organizing for effort, clear expectations, fair and credible evaluations (of student work), recognition of accomplishment, academic rigor in a thinking curriculum, accountable talk (evidence-based discourse appropriate to a discipline), socializing intelligence (intelligence as learned habits of mind), self-management of learning, and learning as apprenticeship.

policies already in place in one or more districts. From that perspective, the SCALE theory of action contemplated external partners supporting district policies in addition to the districts making the policies work more effectively in service of the goal of instructional change. And the district case studies necessarily looked at the entire policy package, not just the part influenced by SCALE (if there was any such distinct part).

As can be seen, the SCALE theory of action for the partnership was emergent and grew along with the equally emergent theory of the Goal 2 immersion projects. The SCALE partnership got off to a fast start because of its superb access to district decision makers and was sustained by the hard work of university and district staff and their openness to the productive working relationships that grew over time.

Summary of the SCALE Theory of Action for District Change

Implementing the SCALE theory of action was an ambitious goal. The reforms were expected to change district policies in a such a way as to deliver strong support for instruction, producing both horizontal and vertical coherence: *horizontal coherence* across policies and *vertical coherence* through the levels of administration between the district and schools. The successful district would have a fully integrated system of distributed leadership and function as a learning organization (Resnick & Glennan, 2002, McLaughlin & Talbert, 2002). The challenge was unavoidable for districts committed to deep and broad instructional improvement. The trick was to find a solution simple and strong enough to do the job.

The District Case Studies

The district case studies[7] were designed to answer two research questions derived from the theory of action:

1. How well was the SCALE theory of action implemented in each district at the level of policy and organization?

2. Did the policies as implemented have the intended effects on teachers and schools?

A set of “panoramic” studies, referenced and heavily relied on in the discussion below, addressed the first research question. Data for these studies was collected in 2003–05. The panoramic studies looked at all policies of instructional guidance (e.g., curriculum, professional development, monitoring, accountability) to see where SCALE had an influence and how different policies might affect each other, then focused on one or two areas in which SCALE had the greatest influence. Addressing the second research question, two in-depth studies looked at selected areas in two districts to see how the policies played out in schools and among teachers.

[7] Evaluation of SCALE was internal, conducted by the Research and Evaluation Team, and had two major parts in addition to the district case studies. Upstream from the case studies, the Building a Partnership Team evaluated the partnership. Downstream from the case studies, broad-scale outcomes like student achievement were evaluated by the Indicators Team, which also conducted targeted studies of particular interventions. The in-depth district case studies overlapped with the targeted studies in evaluating the outcomes of selected interventions, like sixth-grade science in LAUSD.

These two studies were science immersion in LAUSD and interim assessments in Providence. LAUSD implemented science immersion on a broad scale, with major investment from the partnership. Providence was selected as the site to study interim assessments not because of the extent of SCALE influence on that particular policy, but because of the value of having one in-depth study focus on mathematics, the timing (implementation during the SCALE grant), and important issues uncovered by earlier work on the impact of interim assessments in LAUSD.

Research methods for the panoramic studies were ethnographic in the sense that respondents and other data sources (e.g., documents) were chosen according to how well they answered the guiding research questions in the district context. Semistructured interviews started with district leaders and continued until the research questions could be answered, using prior interviews in part to identify key respondents for later interviews (respondents who played some key part in the story) and in part to re-interview respondents about key facts. Fact gathering continued until analytic closure was reached. Different clusters of respondents were involved for mathematics and science at the middle management level (under common leadership at the top). Interviewing began with the departments concerned with curriculum and instruction but gradually expanded to reach other departments and layers—for example, those concerned with supervision of school principals and the local superintendents (in the two large districts). Interviews were coded in a qualitative database, with results analyzed and discussed in multiple meetings of the project team.

Interviews (and focus groups) were an important part of the in-depth studies as well. A survey was used in the Providence study of interim assessments. The considerably larger and more ambitious in-depth study of science immersion in LAUSD included standardized measurements of teacher knowledge, protocols for classroom observation, and district-administered tests of student achievement.

The summaries of the case studies offered below focus on the one or two areas of policy in which SCALE had the greatest influence and for which the most extensive data was collected (as explained above).

Denver Public Schools: Strong Model, Low Funding, Loose Coupling

Denver Public Schools was a good example of how to implement the SCALE theory of action in a large district. The district's chief academic officer was the liaison to SCALE, a leading fellow of IFL, a prime architect of the IFL-based SCALE theory of action, and the architect of the closely parallel DPS theory of action. That configuration of overlapping roles (examined in the case study) ended in 2005 when Denver appointed a new superintendent, who in turn appointed a new chief academic officer. DPS continued as a SCALE partner under a new strategic plan adopted by the district, with SCALE supporting specific math-science initiatives. But the original experiment lasted long enough to reveal important strengths and key structural problems.

The case study tracked the Denver experience from the beginning of the SCALE partnership to a point in 2005 just before the change in district administration (Clifford & Mast, 2005). Two related aspects of the reform were examined: (a) an effort described as *horizontal integration* aimed at fostering a shared vision of good instruction across district and regional leadership; and (b) a second effort described as *vertical integration* (or tightening of governance) aimed at extending instructional leadership of the central office to the field level, among teachers and schools. Together, this kind of horizontal and vertical integration can be understood as an exercise in distributed leadership (getting key actors across the district on the same page about quality instruction; Halverson, 2004), as can the SCALE theory of action itself.

The tools (supportive policies and practices) of horizontal integration included professional development of administrators and teacher leaders in the IFL approach, including the principles of learning, monthly meetings of a cross-departmental education cabinet, and selection of external partners and support based on the principles of learning and the district theory of action.

The tools of vertical integration included creating the office of chief academic officer as an assistant superintendent, creating quadrants to bring administration closer to schools, reorganizing the central office into subject matter specialties (replacing the elementary/secondary division), adding personnel such as division directors and curriculum coordinators, and implementing a curriculum redesign process for selecting, supporting, field-testing, and going to scale with new curricula.

Every aspect of horizontal integration succeeded to some extent. Upper level administrators viewed IFL professional development as important in establishing a common vision for teaching and learning. Administrators from all levels of the system met and exchanged views in weekly meetings of the education cabinet. The common vision of teaching and learning and the authority of the chief academic officer led to an educationally coherent strategy for acquiring resources and joining partnerships. The chief academic officer told the SCALE partnership that the district used “laser beam” criteria for screening external resources and evaluating their conformity with the district program. SCALE itself was accepted because of its integration with IFL and its ability to offer valuable financial support for new math and science initiatives (mostly for professional development).

Vertical integration also succeeded in several ways. Six new curriculum packages were selected under the curriculum redesign process (most before SCALE): *Everyday Mathematics* (elementary level), *Connected Mathematics* (middle level), *Cognitive Tutor Algebra* (secondary level), *Discovering Geometry: An Investigative Approach* (secondary level), *BSCS Biology: A Human Approach* (secondary level), and *BSCS Science Tracks* (fifth-grade science). During the SCALE period, the district was engaged in city-wide implementation of *Connected Mathematics*, which got as far as offering extensive professional development for teachers, online best practice guides, three programs for teacher leaders and specialists following the cascade (train-the-trainer) model, and programs for central office administrators. Field-test teachers met with central office staff weekly to address weak points in the curriculum, professional development needs, and supplementary material. The chief academic officer believed that professional development was the key element of long-run support from the central office to the schools, and central office administrators said that involving teacher leaders and teachers in all steps of redesign helped build ownership and support of the resultant designs, while providing professional learning experiences.

Although horizontal and vertical integration had some success, the data also showed powerful constraints on both efforts. Regarding horizontal integration, although curriculum and instruction (C&I) middle managers attended cabinet meetings with managers of other departments, "loose coupling"—with its typical pattern of limited communication and coordination—prevailed outside the meetings. C&I staff were schooled in the IFL principles of learning because they had been personally trained in them, but managers from other departments were not. School principals attended the cabinet meetings where C&I-sponsored professional development was discussed, and C&I managers were aware of the separate program in the district for professional development of school principals, but the two kinds of professional development were not coordinated in spite of the potentially important role of school principals as instructional leaders.

Vertical integration had its own set of constraints, of which limited coaching resources was the most powerful. Literacy had 135 full-time coaches for 731 K–12 teachers (a 1:7 ratio). Math had 30 part-time coaches who traveled from school to school serving 402 teachers (a 1:14 ratio, with limited hours available). Science had no coaches. Literacy coaches were trained in the IFL coaching method. Math coaches received less systematic training, had limited time to spend with each teacher, and at the school level sometimes were pressed to provide materials and basic support. In science, lacking both coaches and quadrant specialists, C&I administrators recruited and trained teacher leaders who had full-time teaching loads. Vertical integration was also constrained by the lack of line authority of C&I over quadrant math curriculum specialists (who served as math coaches). C&I worked around the problem by hosting a weekly math reading group. Lacking quadrant specialists, the science division worked with teachers through curriculum adoption and redesign committees.

In conclusion, mathematics and science reform in DPS was intelligently designed and partially executed according to the IFL/SCALE theory of action expressed in DPS as horizontal and vertical alignment. The curriculum redesign process did succeed in structuring professional development of teachers around textbooks selected to fit a unified and deeply grounded vision of learning animating the central office. But horizontal integration was limited by loose coupling and insulation of departments from each other, especially the separation of teaching and learning from the supervision of school principals, which left school management outside the circle of distributed instructional leadership. Vertical integration suffered from limited resources for coaching of teachers. To a significant extent, the C&I department in the central office did not succeed in extending its influence across departments or pushing the instructional vision out to schools, despite the favorable influence of the chief academic officer.

Madison Metropolitan School District: Pushing Science Standards out to Schools, Partial Alignment with District-Wide Policies

MMSD was one of the two midsized districts in SCALE (the other was PPSD) and the only district that was not a member of IFL when it joined the partnership. But the district's pre-SCALE reform efforts were consistent with the SCALE theory of action and became better aligned over time, relying strongly on partnership resources. MMSD did join IFL as an affiliate for one year early in the partnership, and IFL left a distinct imprint, for example, in the incorporation of the principles of learning in various aspects of instructional guidance and professional development. In many ways, UW-Madison played the role of IFL for MMSD, and the historically strong relationship of the university with the district in mathematics and science became stronger and more coherent in ways congruent with the SCALE theory of action (Hora & Millar, 2007).

The major initiatives designed to improve math and science instructional guidance were undertaken within the MMSD Department of Teaching and Learning (T&L), which had responsibility for the main elements of the SCALE theory of action: standards, curriculum, and instructional practices; professional development of teachers; monitoring; and assessment. The district staff supported a shift to inquiry-based curricula at all levels because they believed that these curricula provided teachers with materials and pedagogical approaches that were both rigorous and responsive to the needs of all students. Curricular coherence was tightened through implementation of a standards-based student reporting system at the elementary level and a proposed extension of the system to the middle school level.

Elementary and middle school science was a good example of increased curricular coherence. SCALE was involved with a series of efforts to upgrade the quality of the FOSS science curriculum that the district had adopted in the late 1990s. These efforts included fielding a survey of teachers' instructional practice, analyzing state student achievement data, conducting a survey of the enacted curriculum, and revising the scope-and-sequence document. The scope-and-sequence document was used to align and coordinate the major dimensions of instructional guidance specified in the SCALE theory of action (e.g., the "big ideas" to be emphasized in the curriculum and developed across grade levels; professional development offered by instructional resource teachers; the standards-based report cards and grading guides).

District-level content specialists in math and science were increasingly using professional development models that were ongoing, collaborative, on-site, and embedded in the daily work of teachers. Instructional resource teachers conducted professional development sessions and worked on-site with grade-level teams, individual teachers, learning coordinators, and school principals, playing the same role as coaches in the IFL model and using some of the IFL approach. Early in SCALE, some instructional resource teachers attended IFL sessions on the principles of learning, became enthusiastic about the principles, and incorporated them in their own work (for example, the principle of organizing for effort).

The partnership played a multifaceted role in these activities that is only partially summarized here. MMSD used SCALE monies to fund the salaries of two instructional resource teachers. A middle school math instructional resource teacher doubled the capacity of the district to provide professional development to middle school teachers using *Connected Mathematics*. A secondary science instructional resource teacher doubled the capacity (from one to two) of the science staff across all grades. SCALE partners from UW-Madison participated actively in the district's K–8 Science Scope and Sequence Review Committee. SCALE Research and Evaluation Team members facilitated discussions of data from the state assessment, the science survey, and promising professional development practices. SCALE staff from the UW-Madison Center for Biology Education provided valuable insights into science inquiry and ongoing professional development efforts with teachers and schools. And the SCALE Immersion Design Team provided MMSD science staff and teachers with ongoing assistance in the development of immersion units in the district.

The biggest challenge for the SCALE theory of action was a lack of obvious coherence and coupling between the instructional initiatives just described and a series of ambitious, equity-oriented district-wide initiatives that transpired at the same time. Regarding the latter, work continued on a set of three student performance policies adopted by the board of education in the late 1990s (all students proficient in reading by third grade; algebra and geometry completed by the 9th and 10th grades, respectively; and 94% attendance throughout the district). The district also adopted an education framework built around engagement, learning, and relationships, viewed through the lens of race, and incorporating a school-based team model for bringing professional learning communities to scale in all schools. The district launched a district-wide site-based system for differentiated instruction based on an assessment of student needs and corresponding triage of learning options. There was also a relatively new school improvement planning process commencing with a year of needs assessment followed by 4 years of implementation. Schools were asked to orient their goals to the board of education priorities and the district's educational framework. None of these initiatives was managed by T&L.

Opportunities for coordination of instructional guidance managed by T&L with these initiatives did exist and were utilized to an extent not fully investigated by the case study. One venue for collaboration was the superintendent's management team. The instructional council, a subgroup of the management team, focused on teaching and learning. Professional development for school principals was managed by the assistant superintendents and focused on high-quality teaching and learning, as well as the district-wide initiatives and other common needs. School improvement plans might focus on mathematics and science instruction and involve cooperation with T&L, but only if a school's needs assessment identified those areas as priorities. Some cross-departmental integration and collaboration did occur. IFL's principles of learning were selectively incorporated into the educational framework and its implementation. Contact between T&L and school principals was sometimes active, as in this example:

> One elementary [instructional resource teacher] has been running a regular meeting for the six principals in the schools with which she works. The discussions with principals focus on such issues as how math education is changing, the development of students' math content knowledge, assessment of student understanding, and the kinds of support teachers need to collaborate around math instruction. (Scholl, 2005, Elementary Math section)

But in the absence of explicit and planned connections, one can also see the difference between MMSD and a fully integrated IFL district like PPSD (discussed next), where school principals received extensive professional development from IFL. The MMSD case study concluded:

> While SCALE resources and influence have been leveraged largely within the Teaching and Learning Department and "close to the ground," their impact could be significantly heightened by the changes underway at the system level. Thus far, it is unclear whether the system-wide changes and the changes within math and science instructional guidance will come together to create a fully-aligned and coherent approach that is visible from the top-level district administration through to the classroom level. (Scholl, 2005, Interrelation Between System-wide Change and Change in Math/Science Instructional Guidance section, last ¶)

Providence Public Schools: Limited Success with Interim Assessments

Prior to joining SCALE, PPSD had become something of a model IFL district. Central office staff and school principals received IFL training. Coaches trained by IFL were placed in every elementary and middle school. School principals were trained and conducted “learning walks,” using the IFL method of visiting teachers’ classes, giving feedback, and providing support. Coaches were trained in *content-focused coaching* (elementary schools) and *disciplinary literacy* (middle and high schools), both conforming to the IFL principles of learning. PPSD was also in the midst of rolling out and implementing a series of inquiry-based courses, including *Investigations* (elementary school math), *Connected Mathematics* (middle), FOSS (middle school science), and *Physics First* (high school). For all of these reasons, the district was generally a good fit with SCALE and its theory of action.

In the early days of the SCALE partnership, IFL and other SCALE partners worked cooperatively on a range of projects (White, 2005). At the top leadership level, a full-day meeting was held in November 2004, at which Lauren Resnick and the PPSD management team discussed ongoing efforts to implement a rigorous curriculum in mathematics and science. With SCALE funding, IFL coached PPSD staff in the design and implementation of nine after-school professional development sessions in mathematics for middle and high school teachers. IFL worked with mathematics coaches on teachers’ ability to select tasks with a high level of cognitive demand, after which the team visited teachers’ classrooms, looking at instructional quality and feedback on students’ work. Major partnership assistance focused on implementation of FOSS, which the district had put on the fast track for development of the science curriculum. PPSD teachers and administrators attended a FOSS workshop in Madison. The district used SCALE funds to appoint a full-time consultant on science curriculum development and contracted for FOSS training and science literacy professional development with the East Bay Educational Collaborative.

Two major constraints on the SCALE theory of action were turnover in district personnel and policy incoherence produced by different models of professional development and testing. As for turnover, during the SCALE grant, the district had three superintendents and one interim superintendent; numerous shifts in central office personnel, with several positions left vacant for months at a time; and rapid turnover among school principals (who had been trained by IFL). Fortunately for SCALE, a core of district administrators strongly supportive of the SCALE vision stayed in place during the partnership period. But the turnover took its toll.

A good example of the policy dissonance produced by having a variety of providers of professional development was the Math Matters program for elementary and middle school mathematics teachers, a standards-based professional development program created by WestEd with the support of NSF and the California Department of Education. Tom Lester, a nationally known mathematics trainer from WestEd, provided training for coaches through the Teacher Leadership Academy Program. The size of the Math Matters program was substantial. For example, by the 2004–05 school year, PPSD had employed 27 mathematics coaches (approximately one at each elementary school) applying Math Matters to the mathematics *Investigations* curriculum. District and school staff did succeed to some extent in reconciling different approaches to professional development—for example, by infusing Math Matters with the principles of learning. But respondents still identified a disconnect between Math Matters and the IFL initiative—for example, in the greater emphasis in Math Matters on skills and discrete problem solving.

The coexistence of the IFL program and PPSD's major push to implement interim assessments of student achievement provides another example of policy incoherence in the district. We decided to conduct an in-depth study of the implementation of these assessments (Clune & White, 2008) because of questions about policy alignment raised by a prior study of interim assessments in the Los Angeles district (discussed in the section on LAUSD). PPSD initiated the program during the 2004–05 school year and terminated it in the fall of 2007.

The district initially rolled out the interim assessments in Grades 2–8 in English language arts and mathematics. Teachers on the committee and teacher coaches reviewed the results to identify gaps. As a result of the committee's work, approximately 2,000 test items were compiled in an item bank maintained by the district assessment office. Each grade-level interim assessment was composed of 21 questions—20 multiple-choice and one open-ended. The items were keyed to grade-level expectations, with each expectation covered in at least one test item. In mathematics, the open-ended question could be a multi-step problem with three or four parts that required students to give the correct answer and show their work. Results on the multiple-choice items were reported by the district, while teachers scored the open-ended question for their own students using a rubric.

District officials articulated three goals for the interim assessments: (a) completing what the district referred to as a "great alignment" of the PPSD scope and sequence, grade-level expectations (or for high schools, grade span expectations), curriculum, and assessments; (b) providing practice and preparation for the new state assessment, the New England Common Assessment Program (NECAP); and (c) providing data for teachers on the instructional needs of students, present and future. The case study surveys of teachers and administrators at both district and school levels revealed that the district met its three goals at least to some extent. Teachers aligned instruction with the assessments (e.g., filling in gaps in coverage), believed that students were better prepared for the state test, and used assessment data to meet the instructional needs of current students. School support and leadership were viewed as essential to best practice. Site-based professional development, such as planning time and assistance from coaches, was mentioned frequently as an important venue for discussing the assessments, especially in elementary schools. Our earlier panoramic study of the district had shown that practices of site-based professional development had been built up during prior waves of reform (White, 2005). For example, "daily twos" and "head problems"—short exercises that were the most commonly mentioned way of reviewing the assessments with current students—were established by the Math Matters program of professional development.

At the same time, it was unclear whether the interim assessments would be capable of producing the major gains in student achievement attributed to formative classroom assessment (Black & William, 1998; Stiggins, 2002; Brookhart, 2005). Effective formative classroom assessment normally depends on deep professional development for teachers and schools in the skills of using student work to orient instruction and motivate students. Although the case study showed that some teachers were reviewing the assessments with current students, the practice was limited. Teachers still expressed problems with integrating results received 2 weeks after assessment; many of the instructional changes consisted of adjusting content coverage for subsequent years rather than working with current students; district professional development seemed directed at interpreting data rather than improving instruction, and teachers expressed the need for more professional development; and school support seemed highly variable—not part of the design or implementation of the interim assessments project—and subject to erosion from budget cuts (affecting, for example, the number of coaches).

It is doubtful that that the interim assessments were fully consistent with the SCALE theory of action. The IFL model stressed professional development of coaches and teachers and deep learning of the subject matter rather than alignment of the enacted curriculum with detailed grade-level guides and items on a standardized assessment. Moreover, professional development oriented to classroom instruction was the weakest part of the interim assessments. The two models running in parallel were not necessarily completely inconsistent, but they clearly were not designed to operate as a coherent package. They are best understood as examples of fragmented rather than coherent reform (Cohen, 1982, Cohen & Spillane, 1992; Spillane, 1998).

Los Angeles Unified School District: The Math Planning Group and Science Immersion Units

LAUSD—with its enrollment of some 750,000 students and its 35,000 teachers and 600 math coaches—was the giant of the SCALE partners. If the four partner districts were thought of as Russian dolls, Denver would fit inside one of the 11 local districts of LAUSD, while Madison and Providence would each fit in one of the four quadrants of the Denver district. The great size of LAUSD, its complexity, and the extra layers between its central office and schools posed a super-sized challenge for the SCALE goals of horizontal and vertical alignment (Osthoff, 2006). At the same time, the sophistication of the central office staff, together with the goal-oriented approach of then Superintendent Roy Romer, allowed for a high degree of innovation and follow-through.

LAUSD was a member of IFL when it joined the SCALE partnership, and IFL Director Lauren Resnick had a collaborative relationship with Superintendent Romer. SCALE became active in all areas of its theory of action. For example, SCALE had significant involvement in the efforts of the LAUSD Math Planning Group (MPG) to upgrade instruction in mathematics (Osthoff, 2006). The MPG consisted of administrators and mathematics instructional support staff from the central office and local districts charged with implementing instructional guidance policies. One modification considered by the group was the addition of model lessons that would guide teachers on how to set up, explore, share, discuss, and analyze challenging everyday problems. Model lessons were to be disseminated and supported through a new strategy for expedited teacher training concentrating initially on helping a small group of coaches gain proficiency with the lessons. These coaches would in turn work with teachers at their schools in groups and in individual classrooms to implement the model lessons.

The MPG also looked at newly implemented periodic assessments. The assessment scores were very low, especially at the Algebra I level, motivating the MPG to look for ways of improving district support for teaching and learning of standards-based content of the kind emphasized on district and state assessments. As part of the MPG, the SCALE Research and Evaluation Team, in cooperation with the LAUSD Program Evaluation and Research Branch, analyzed data from focus groups to determine how teachers viewed the interim assessments and other elements of instructional guidance (Osthoff & Cantrell, 2004). The study, which clearly influenced district leaders, found that many teachers thought that the curriculum had become overcrowded. Most saw textbooks as the primary source of instructional guidance and therefore experienced other guidance (like the assessments) as calling for instruction on top of what was covered in the textbook. Periodic assessments were seen as asking teachers to provide extensive remedial instruction to students, requiring them to suspend instruction on regular course content to address gaps in requisite student knowledge. Teachers also felt burdened by instructional guides that recommended significant use of supplementary materials. Model lessons were seen as nice ideas that could not possibly be squeezed into an already overcrowded curriculum. Thus, district leaders learned that, in this case, what seemed coherent to district managers did not seem so to teachers.[8]

Science immersion units were a strong and sustained intervention of the SCALE partnership in LAUSD. At the heart of immersion unit design is a commitment to engaging students in a full cycle of inquiry (e.g., typically spanning a week or more of science instructional time) into real-life phenomena (e.g., organic decomposition and photosynthesis; electric circuits and magnets). The inquiry approach encourages students to acquire factual knowledge and facility with procedural knowledge (e.g., controlled experimentation, scientific observation) in the context of exploring relationships among important phenomena. This stands in contrast to traditional science instruction, which often treats memorization and procedural knowledge as ends in themselves, paying little attention to conceptual understanding or application of knowledge to real-life scientific questions.

During the SCALE partnership, five immersion units were developed and implemented: *Rot It Right* (transfer and cycling of matter and energy during decomposition; Grade 4), *Weather* (Grade 5), *Plate Tectonics* (Grade 6), *Variation and Natural Selection* (Grade 7), and *Density and Buoyancy* (Grade 8). Implementation of the immersion units was an evolving process. Early on, the partnership adopted a statement melding principles of scientific immersion and the IFL principles of learning (Schunn et al., 2006). From the outset, the UW-Madison immersion team worked with science experts at the level of LAUSD middle management who were close to the point of delivery. The work was enabled by the superintendent, who gave science a high priority; the leadership of top central office managers; and a preexisting culture favoring science inquiry among district middle managers. Early in the implementation period, the UW team assisted with the incorporation of immersion into district instructional guides and professional development for teachers. Complete curriculum units were developed and supported by parallel professional development, through the process of co-construction and co-implementation. By the third year, explicit models were designed for curriculum development, professional development, and facilitation of professional development. All of this was immensely fortified both in planning and implementation by a large U.S. Department of Education grant awarded to SCALE partner California State University, Dominguez Hills (CSUDH). Through this grant, CSUDH and the other California state universities in the SCALE partnership provided professional development for teachers who would implement the immersion units, thereby operating as a local delivery platform for science immersion in LAUSD.

[8] Considerable work was subsequently done by IFL and LAUSD—without SCALE involvement—on model lessons and related issues. This work is not reflected in the SCALE district case studies or reported in this paper.

A substantial in-depth case study was undertaken of the Grade 6 unit on plate tectonics (Osthoff et al., 2007). Multiple instruments were used to measure the quality of classroom instruction and student learning. The study is still in progress, but initial findings indicate that (a) teacher content knowledge was improved through professional development; (b) fidelity of implementation was excellent, with most teachers attempting to teach lessons in order and in their entirety (though the typical teacher was able to complete only about half the unit in the time allotted); (c) student engagement in immersion classrooms was high; and (d) the instruction of immersion teachers shifted toward higher levels of cognitive demand (such as conceptual understanding). No positive effects on student achievement (as measured by periodic assessments) were observed, due in part to the difficulty teachers had completing the immersion units when teaching them for the first time. Implementation challenges were observed in three areas: (a) pacing (covering the entire unit), (b) aspects of instructional quality, such as skill with open-ended questions and use of rubrics for scoring student work, and (c) capacity building, especially involvement of school principals and other administrators.

Two limitations were professional development and alignment with the periodic assessments. Professional development moved closer to the school in that it was offered for the district science specialists who then provided it to teachers at a centralized location. But professional development reached only a minority of science teachers overall—rarely all teachers in a given school—and it was not available at the school level. In other words, LAUSD did not find a way to push the professional development all the way down to the school for more than a few teachers per school in a subset of schools. Regarding alignment, the immersion unit placed far greater emphasis on some cognitive demand categories (*communicate understanding*, *analyze*, *make connections*) than the periodic assessments, which focused heavily on memorization and procedural knowledge (with some attention to *communicate understanding*). Consequently, the more successful teachers were in implementing the immersion units with fidelity, the more the alignment of the units with the periodic assessments suffered.

In summary, the sixth-grade science immersion units were implemented with a considerable degree of coherence across the curriculum and professional development. Monitoring and accountability were provided externally through the SCALE in-depth case study and internally through the interim assessments and efforts of district staff. Two limitations were the absence of professional development at the school level and lack of alignment with the interim assessments.

Conclusion

The SCALE theory of action was successful in many ways. Excellent access to top management allowed the partnership to assist with multiple dimensions of instructional guidance. Major reforms unfolded with partnership assistance in every district: curriculum redesign and professional development in DPS, new science standards and site-based professional development in MMSD, professional development of math coaches in PPSD, and model lessons and science immersion in LAUSD. Horizontal and vertical alignment was a central focus, each reform containing multiple elements of instructional guidance coordinated with each other and a strong focus on support for schools and teachers. Sustainability can be difficult to recognize so close to the events (Century & Levy, 2002), but the district reforms assisted by the partnership had many of the elements of a sustainable impact: a unifying vision of instruction, seemingly durable changes in district policy and organization, and leadership, as well as apparently secure institutionalization of curricula, professional development, monitoring, and accountability.

The biggest obstacles were turnover in district leadership, loose coupling of the multiple elements of system-wide reform, insufficient budget for adequate school site support (e.g., coaches), and lack of horizontal and vertical alignment. Horizontal alignment was never complete due to parallel systems of leadership within districts, professional development, and assessments. Vertical alignment also was constrained by loosely coupled parallel systems of school improvement and school principal training, insufficient numbers and training of coaches, and the complexity of integrating multiple streams of instructional guidance for teachers. In many ways, the obstacles were the familiar problems of fragmentation and complexity found in other studies of systemic reform (Cohen, 1982; Cohen & Spillane, 1992). In the obstacles, one can also see a hole in the SCALE theory of action noticed early in the partnership by Halverson (2004): the absence of a clearly articulated link between policy at the top and schools.

One can also see progress in the theory of action over time in lessons learned. From the NSF point of view, organizational learning through experimentation is a main goal of the math and science partnerships, which are aimed more at discovery or exploration than implementation through "exploitation" of an existing best practice (Axelrod & Cohen, 2000). From the perspective of horizontal and vertical integration, science immersion represented the most coherence, with tight packaging of a detailed curriculum, professional development for teacher trainers, and professional training for teachers quite close to the school level. Late in the partnership, SCALE assisted with implementation of a similar package in LAUSD, Agile Mind Biology (Charles A. Dana Center, 2008). Agile Mind (which was not the subject of a case study) is a package of highly articulated classroom lessons and formative assessments supported through initial face-to-face training and daily lesson-by-lesson online professional development of teachers. Students also have access to the system online. Usage and sophistication of instruction are monitored at Agile Mind headquarters, with feedback provided to sites. Affordability is insured by strict cost controls on fees and dial-up computer access for teachers and students. Fundamentally, these interventions move instructional support much closer to the school and classroom, thereby gaining both power and coherence at the point of delivery. Instructional systems adopted with a high level of buy-in at the school level, as required by Agile Mind, seem more likely to survive changes in top leadership.

The idea of a coherent package of instructional guidance delivered directly to schools can be seen in a broad class of interventions—for example, replacement units in mathematics implemented in California (Cohen & Hill, 2000) and RtI (*response-to-intervention* or *response-to-instruction*) models designed especially for differentiated instruction (Foorman, Kalinowski, & Sexton, 2007). A familiar example is Robert Slavin's highly successful intervention Success for All (Borman et al., 2007). Success for All is a whole-school program aimed at the acquisition of basic reading skills by all children; the program has been implemented in more than 1,200 mostly high-poverty Title I schools in 46 states. An excellent example of "exploitation" of a best practice, Success for All is purchased as a comprehensive package, including materials, training, ongoing professional development, benchmark assessments, and a plan for organizing resources. Students in Success for All schools spend most of their day in traditional age-grouped classes, but are regrouped across grades for reading lessons targeted to specific performance levels.

Teachers assess each student's reading performance at 8-week intervals and make regrouping changes based on the results. Most students who need additional help receive one-to-one tutoring to get them back on track. Cooperative learning is used to build academic skills. Each Success for All school has a full-time program facilitator who oversees the daily operation of the program, provides assistance when needed, and coordinates the various components. Like Agile Mind, Success for All has been designed to operate affordably within existing school budgets

All the elements of the SCALE theory of action are present in Success for All at the school level: the goals of high-quality instruction for all students leading to high student achievement, expertise and support provided by an external partner, a specified curriculum, site-based professional development, monitoring and adaptation of instruction, organizational restructuring, and managerial training. The essential point here is the underlying principle rather than the particulars of specific programs—namely, that the challenges to horizontal and vertical coherence uncovered in the SCALE district case studies are best met by providing a complete package of instructional guidance fashioned for coherent operation at the point of delivery, rather than by expecting elements or pieces of instructional guidance created at the district level to come together on their way to schools. The grand challenge raised by the SCALE district case studies is whether such school models can be successfully replicated across entire districts through some combination of central and school management. Answers to those questions will require an extension of the organizational learning begun in SCALE. According to Peter Senge's 1990 book on systemic thinking in business settings, *The Fifth Discipline*, the most powerful organizational learning occurs through the leverage achieved by replacing a negative dynamic (such as vertical and horizontal incoherence) with a positive one (such as a combination of district and school-based models). Through such organizational learning, obstacles and opportunities become better understood, improved practices are designed and adopted, and problems become more tractable (Resnick, Besterfield-Sacre, Mehalik, Sherer, & Halverson, 2007; Halverson, 2006).

References

Axelrod, R., & Cohen, M. (2000). *Harnessing complexity, organizational implications of a scientific frontier*. New York: Basic Books.

Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–148.

Borman, G., Slavin R., Cheung, A., Chamberlain, A., Madden, N., & Chambers, B. (2007). Final reading outcomes of the national randomized field trial of Success for All. *American Educational Research Journal, 44*(3), 701–731.

Brookhart, S. M. (2005, April). *Research on formative classroom assessment.* Paper presented at the annual meeting of the American Educational Research Association, Montréal, Québec, Canada.

Century, J. R., & Levy A. J. (2002). *Researching the sustainability of reform*. Newton, MA: Education Development Center.

Charles A. Dana Center. (2008, October). *Agile Mind Biology: The use of multimedia resources in the high school science classroom*. Presentation at the Mathematics and Science Higher Education Conference, Austin, TX. Retrieved February 24, 2009, from http://www.utdanacenter.org/downloads/octconf/2008/kirkpatrick.pdf

Clifford, M., & Mast, G. (2005). *SCALE panoramic case study: Denver Public Schools*. Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved February 24, 2009, from http://www.scalemsp.org/index.php?q=SCALE_Panoramic_Case_Study_DPS

Clune, W. H. (2005, April). *Case study evaluation design*. Paper presented at the annual meeting of the American Educational Research Association, Montréal, Québec, Canada. Retrieved February 24, 2009, from http://www.scalemsp.org/files/research/Products/Clune_CaseStudy_AERA.pdf

Clune, W. H., & White, P. A. (2008). *Policy effectiveness of interim assessments in Providence Public Schools* (WCER Working Paper No. 2008-10). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved February 24, 2009, from http://www.wcer.wisc.edu/publications/workingPapers/Working_Paper_No_2008_10.php

Cohen, D. (1982). Policy and organization: The impact of state and federal educational policy on school governance. *Harvard Educational Review, 52*(4), 474–499.

Cohen, D., & Hill, H. (2000). Instructional policy and classroom performance: The mathematics reform in California. *Teachers College Record, 102*(2), 294–343.

Cohen, D., & Spillane, J. (1992). The relations between governance and instruction. *Review of Research in Education, 18*(1), 3–49.

Foorman, B. R., Kalinowski, S. J., & Sexton, W. L. (2007). Standards-based educational reform is one important step toward reducing the achievement gap. In A. Gamoran (Ed.), *Standards-based reform and the poverty gap: Lessons from No Child Left Behind* (pp. 17–44). Washington, DC: Brookings Institution.

Halverson, R. (2004). *A distributed leadership perspective on the SCALE theory of action.* Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved February 24, 2009, from http://www.scalemsp.org/files/research/Products/Halverson_DistributedLeadershipPerspective.pdf

Halverson, R. (2006). *A distributed leadership perspective on how leaders use artifacts to create professional community in schools* (WCER Working Paper No. 2006-4). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved February 24, 2009, from http://www.wcer.wisc.edu/publications/workingPapers/WorkingPaperNo.2006-4.php

Hora, M. T., & Millar, S. B. (2007). *A preliminary case study of SCALE activities at the University of Wisconsin–Madison: Factors influencing change initiatives in STEM undergraduate education, teacher training, and partnerships with K–12 districts* (WCER Working Paper No. 2007-2). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved February 24, 2009, from http://www.wcer.wisc.edu/Publications/workingPapers/Working_Paper_No_2007_02.php

Institute for Learning. (2007). *Principles of learning*. Retrieved February 25, 2009, from http://ifl.lrdc.pitt.edu/ifl/index.php?section=pol

McLaughlin, M. W., & Talbert, J. E. (2002). Reforming districts. In A. M. Hightower, M. S. Knapp, J. A. Marsh, & M. W. McLaughlin (Eds.), *School districts and instructional renewal* (pp. 173–192). New York: Teachers College Press.

Osthoff, E. (2005). SCALE *Research and Evaluation Team case study: Los Angeles Unified School District.* Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved February 24, 2009, from http://www.scalemsp.org/index.php?q=SCALE_RET_Case_Study_LAUSD

Osthoff, E. (2006). *Moving From a policy- to system-design focus to increase rigor and equity in teaching and learning: Mathematics instructional reform in Los Angeles Unified School District during the SCALE initiative*. Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research.

Osthoff, E., & Cantrell, S. (1994). *LAUSD mathematics teacher and coach focus groups: Views of district instructional guidance from the field* (SCALE Technical Report). Retrieved February 24, 2009, from http://www.scalemsp.org/index.php?q=LAUSD_Groups_Views_from_Field

Osthoff, E., Clune, C., White, P., Ferrare, J., Kretchmar, K., & Kelly, K. (2007). *Tentative findings of the SCALE study of middle school immersion in LAUSD: A conversation with*

LAUSD immersion professional development facilitators. Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved February 24, 2009, from http://www.scalemsp.org/files/research/Products/SCALE_Gr_Immersion_Study_Yr_5_report.pdf

Resnick, L. B. (2006). *The SCALE teaching and monitoring systems: An analytic guide for district planning.* Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research.

Resnick, L. B., Besterfield-Sacre, M., Mehalik, M., Sherer, J. Z., & Halverson, E. (2007). A framework for effective management of school system performance. *Yearbook of the National Society for the Study of Education, 106*(1), 155–185.

Resnick, L. B., & Glennan, T. K. (2002). Leadership for learning: A theory of action for urban school districts. In A. M. Hightower, M. S. Knapp, J. A. Marsh, & M. W. McLaughlin (Eds.), *School districts and instructional renewal* (160–172). New York: Teachers College Press.

SCALE. (2003). *Revised 5-year strategic plan to the National Science Foundation.* Madison: Author.

Scholl, L. (2005), *SCALE case study: Madison Metropolitan School District*. Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved February 24, 2009, from http://www.scalemsp.org/index.php?q=SCALE_Case_Study_MMSD

Schunn, C., Millar, T., & Lauffer, D. (2006). *Immersing all K–12 students in extended inquiry science and design* (SCALE Goal 2 concept paper). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved February 24, 2009, from http://www.scalemsp.org/index.php?q=Immersing_Students_Extended_Inquiry

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization.* New York: Doubleday.

Spillane, J. P. (1998). State policy and the non-monolithic nature of the local school district: Organizational and professional considerations. *American Educational Research Journal, 35*(1), 33–63.

Spillane, J. P., Halverson, R., & Diamond, J. B. (2001). Investigating school leadership practice: A distributed perspective. *Educational Researcher, 30*(3), 23–28.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment FOR learning. *Phi Delta Kappan, 83*(10), 758–765. Retrieved February 24, 2009, from http://www.pdkintl.org/kappan/k0206sti.htm

Weiss, C. (1998). *Evaluation: Methods for studying programs and policies* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

White, P. (2005), *SCALE case study: Providence Public School District*. Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research.