

Formative Feedback Systems and the New Instructional Leadership

Richard Halverson

Department of Educational Leadership & Policy Analysis
Wisconsin Center for Education Research
University of Wisconsin–Madison
halverson@education.wisc.edu

Reid B. Prichett

Department of Educational Leadership & Policy Analysis
Wisconsin Center for Education Research
University of Wisconsin–Madison
rbprichett@wisc.edu

Jeffery G. Watson

Wisconsin Center for Education Research
University of Wisconsin–Madison
jgwatson@wisc.edu



Copyright © 2007 by Richard Halverson, Reid B. Prichett, and Jeffery G. Watson
All rights reserved.

Readers may make verbatim copies of this document for noncommercial purposes by any means, provided that the above copyright notice appears on all copies.

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

Halverson, R., Prichett, R. B., & Watson, J. G. (2007). *Formative Feedback Systems and the New Instructional Leadership* (WCER Working Paper No. 2007-3). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved [e.g., May 15, 2007,] from <http://www.wcer.wisc.edu/publications/workingPapers/papers.php>

The research reported in this paper was supported by the National Science Foundation under Grant # 0347030 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.

Formative Feedback Systems and the New Instructional Leadership

Richard Halverson, Reid B. Prichett, and Jeffery G. Watson

Formative feedback systems are systems of structures, people, and practices that help teachers and administrators translate testing data into practical information for everyday use. Formative feedback systems live at the heart of systemic school improvement efforts. Without accurate and timely information on the results of intended interventions, school leaders and teachers fly blind in their efforts to link what they expect to what actually happens in their schools. In this paper, we argue that school capacity to use data consists largely of the ability to generate and make good use of information on the core processes of teaching and learning. We use the concept of formative feedback systems to describe the socio-technical processes that leaders and teachers coordinate to develop this capacity.

Recent research on data-driven decision making has focused on how (and whether) schools can provide the right kinds of data to help schools meet the demands of high-stakes accountability programs (Mandinach, Honey, & Light, 2005; Murnane, Sharkey, & Boudett, 2005; Thorn, 2001; Wayman, 2005; Watson & Mason, 2003). These researchers have shown that access to information is only one part of the school improvement process. Leaders in successful schools have engaged teachers in collaborative redesign of the school curriculum and have developed methods to determine whether these curricular innovations made a difference in student learning (Smylie, 1994; Hallinger, Bickman, & Davis, 1996; Lachat & Smith, 2005). Without the social, technical, and professional capacity to act effectively on information, however, schools often simply retrench with status quo practices (Spillane, 2000).

The use of student assessment data as a measure of school effectiveness, a central premise of NCLB, requires that schools see how achievement tests result from day-to-day classroom practices. This is a considerable organizational challenge, partially because the traditional loose-coupling structures prevent strong links between classroom and organizational outcomes, and partially because the standardized test results are ill-timed to make a difference in classroom practices.

The press for meeting the demands of high stakes accountability policies has led many schools to reframe the tasks of school leadership. As Richard Elmore (2000) has described, the context of standards-based reforms alters the landscape for instructional leadership. Since school staff cannot rely on standardized test results to inform changes in their classroom-level practices, leaders and teachers who intend to meet accountability demands must engage in a two-level redesign—first to link everyday classroom practices with school-wide outcomes, and second to develop data systems that give teachers much more local, ongoing measures of student learning. Despite Elmore's warning about the lack of capacity in local schools to fundamentally change organizational practices, practitioners across the country are engaged in redesigning their schools to meet high-stakes accountability requirements (Yeh, 2005; Light et al., 2005; Wayman, 2005; Murnane & Sharkey, 2005).

Our recent research into how schools develop the capacity to use data to effectively improve student learning has shown us how local actors develop data-driven instructional systems to improve classroom practice (Halverson, Grigg, Prichett, & Thomas, in press;

Formative Feedback Systems and the New Instructional Leadership

Halverson & Thomas, 2007). One key feature of this research has been uncovering the operation of *formative feedback systems*. Typically, a formative feedback system consists of schools' repositioning a central actor, aligning curriculum, constructing assessments, and developing collaborative discussion and decision-making spaces in order to turn the information generated by their curriculum into useful information. We illustrate the concept of a formative feedback system in action through a detailed reconstruction of the literacy program in one elementary school noted for the school's leadership around data-driven decision making. This example shows how leaders and teachers construct formative feedback systems, and points to design issues schools may face in constructing formative feedback systems of their own.

Formative Feedback Systems: Some Background

Originally a concept developed in cybernetics and engineering research (von Bertalanffy, 1969), organizational feedback became a popular term in 1990s social science and organizational research (Senge, 1994; Richardson, 1991; Greve, 2003). Feedback is generated by a system and is looped back to control system processes. In its simplest form, a feedback system consists of four main parts: signals, sensors, signal processors, and controllers (von Bertalanffy, 1969; Richardson, 1991). Signals contain information from within or outside of the system. Sensors detect the presence of the signals, and processors establish the significance of the signal. Signal processors analyze and interpret the signal meaning, and controllers determine the action to be taken as a result, which may result in a new signal that acts as new input into the information system.

In organizations, feedback information about organizational performance is used to influence the structures that guide organizational behaviors, using negative feedback to create new behaviors or positive feedback to reinforce status quo behaviors (Argyris, 1990). At its core, successful formative feedback about organizational performance is a form of interpersonal communication (Ilgen, Fisher, & Taylor, 1979). This suggests that a consideration of feedback systems must move beyond generating information to describe a *performance information + interaction = action* model. This emphasis on communication brings out the reality of how performance data have traditionally been received in organizations (Ilgen & Davis, 2000; Kluger & DiNisi, 1996). Performance feedback is often seen as information used to punish or harass, and can be seen as irrelevant to the evaluated practices. Such feedback can also be seen as untimely, and related to matters outside the control of practitioners. The processes of collecting performance feedback can be perceived as wasteful by practitioners, at least compared to the potential benefit realized by the information (Kreitner & Kinicki, 2001).

Formative vs. Summative Feedback

One problem in understanding feedback is the distinction between summative and formative feedback information. Summative feedback describes the *results* of processes, while formative feedback is used to *inform* and *adjust* the process as it unfolds. Still, the distinction between summative and formative often lies in the perception of the communicators, not in the information itself. Thus, information generated, for example, through shared assessments or peer observation can be interpreted and used as evidence to summatively judge and discipline teachers, just as standardized test scores can be used to formatively reshape instructional

Formative Feedback Systems and the New Instructional Leadership

practices. Information must be interpreted if it is to be used as feedback that can improve organizational processes (Nadler, 1979).

Formative Feedback in the Classroom

Black and Wiliam's (1998) work highlighted the critical role that formative feedback plays in classroom learning. They described how formative feedback works on three levels in the classroom: the teacher level, the student level, and at the level of teacher-student interaction. At the teacher level, teachers need accurate information about the specific processes and outcomes of student learning to effectively shape their teaching. Students also need accurate self-assessments to guide their learning processes. The formative feedback process comes to life through student-teacher and student-student interaction in the form of questions that highlight learning gaps and through discussions that show how these gaps might be addressed. Black and Wiliam's summary of prior research on formative feedback demonstrates impressive learning gains within the classroom. A formative feedback system perspective suggests that school organizations might benefit from similar learning gains as classrooms when formative feedback practices are extended across schools.

School-Wide Formative Feedback

A school's formative feedback system structures opportunities for teachers and school leaders to (a) learn from organizational performance information and (b) adjust instructional programs and practices accordingly. From a systems perspective, a school is a complex, messy information system that issues many performance signals (Wallace & Pocklington, 2002). Some of these signals are acted upon, some misinterpreted, and others simply ignored. Within schools that have proven able to systemically improve student learning, however, there exist closed subsystems capable of responding to performance information (Halverson, 2003; Burch & Spillane, 2003; Gamoran et al., 2003). Extending Black and Wiliam's (1998) three-level description of a classroom-based system to the school would mean that (a) teachers and school leaders would need accurate information on instructional program success, (b) teachers themselves would need accurate information on their own efforts to teach the instructional program, and (c) legitimate opportunities would be provided for interaction to raise questions and discuss solutions. A feedback system across classrooms would generate information signals that measure student and program performance, sensor functions to detect such signals, processor functions to make sense of information signals, and controllers that could use this new knowledge to adjust the instructional process.

Methods

DDIS Study

Our study of *data-driven instructional systems* (DDIS) (Halverson et al., in press; Halverson & Thomas, 2007) investigated how school leaders and teachers are engaged in using data to redesign their local instructional practices. This paper represents data collected during the initial stages of a 5-year National Science Foundation-funded research project designed to study how leaders create social and technical systems to help teachers use achievement data in their

Formative Feedback Systems and the New Instructional Leadership

instruction. The study was organized around a distributed leadership perspective on analyzing school leadership practice (Spillane, 2006; Spillane, Halverson, & Diamond, 2004).

Distributed leadership focuses on how the key tasks of instructional leadership are *socially distributed* among school professionals, and *situationally distributed* across the artifacts, such as policies, programs, and procedures, that constitute the local context. Halverson's (2003) description of the situational distribution of leadership practices emphasized how leaders used artifacts to shape the "system of practice" that constrained teaching and learning practices in schools. The DDIS study sought to identify the artifacts leaders designed and used to support data-driven instructional practices in their schools.

Site Selection

The study design was intended to document leadership practices, assess their fit with the DDIS we developed, and describe the similarities and differences among schools' instantiations of the DDIS. In order to select successful schools, we consulted with educational leaders at the university, state, and district levels to generate a list of elementary and middle schools known for improving test scores and with leaders who were known for using data well with their teachers. From our initial list, we narrowed our sites to nine rural, urban, and suburban schools recognized for strong data-driven decision making and records of improving student achievement. We gave highest priority to schools with the strongest record of improving student achievement.

For the purpose of this study, we have decided to focus on the formative feedback systems present in one school, Pearson Elementary School, to provide an in-depth illustration of the system organization and operation. The early literacy program in the Pearson Elementary School clearly demonstrated the features of formative feedback systems we observed across our schools. In the conclusion, we use data samples from several of the other schools to illustrate or provide contrast to the central themes highlighted in the Pearson case.

Data Collection

In order to document and describe the DDIS at Pearson, we collected a variety of data including: (a) structured interviews with formal and informal leaders, (b) observations, and (c) relevant documents. Previous research in distributed leadership studies (Spillane et al., 2004; Halverson, 2003, 2004) has identified these methods of data collection as the most useful for documenting the systems of practice at a given school. Once schools were identified, we asked key formal leaders to identify the artifacts they felt were critical for their efforts to improve student learning, and to identify the local teachers and staff most involved with designing and implementing these artifacts. These conversations typically identified five to eight artifacts school leaders and teachers had either designed or imported to improve student learning. Typically, the artifacts changed over time, added new features, or were supplanted by other initiatives. Our research was designed to trace how leaders and teachers assembled the artifact systems, to trace how the resulting social and technical systems functioned and influenced the everyday practice of school teachers and leaders, and to tell the story of what local practitioners learned in the process of developing their capacity for data use.

Formative Feedback Systems and the New Instructional Leadership

Most of our observations concerned faculty meetings and other adult gatherings in the schools. We typically observed classroom teaching practice only when invited by staff to see how the data-driven decisions influenced teaching and learning. We spoke with a selection of other teachers in each school to understand their assessment of the relevance and effects of the artifacts identified by school leaders.

Over the course of one and a half school years, we conducted 107 structured interviews with the DDIS school teachers and the formal and informal leaders. Most of these were one-on-one interviews, but there were 2–3 group interviews of 2–4 teachers and specialist staff members for each school site. We conducted 135 observations of classroom teaching sessions, faculty meetings, professional development sessions, data retreats, and other important events as identified by the staff. These observations ranged from 1 to 3 hours per visit. We also collected a variety of documents from every school, such as school improvement plans, staffing charts, budgetary information, and parent/community handouts. The Pearson case data are a subset of the data set collected from across the nine schools. At Pearson, we interviewed 12 teachers and school leaders, and made 15 observations of classroom instruction, faculty meetings, and other professional meetings.

Data Analysis

The study approach to data analysis involved: (a) constructing an initial theory of data-driven decision making—the DDIS framework—based on prior research on how schools meet the demands of external accountability and (b) using the DDIS framework to relate the practices of school leaders who have established reputations for successful use of data to improve instruction. Our analysis draws on a data set composed of individual school case studies. Relying on organizational and school change literature, we developed a DDIS framework that described the six central functions for how data were used in schools (for a more detailed elaboration of the DDIS framework, see Halverson, Grigg, Prichett, & Thomas, 2007 Halverson et al., in press):

1. Data acquisition: What kinds of data do schools collect and rely upon to inform action?
2. Data reflection: What kinds of processes are structured to help schools make sense of the data and to set instructional goals?
3. Program alignment: How do schools determine the gaps in their existing instructional programs?
4. Program design: What kinds of programs, policies, or procedures do leaders and teachers implement to improve student learning?
5. Formative feedback: How do teachers and leaders know whether the programs they have implemented are making a difference in student learning?
6. Test preparation: How are students prepared for the testing process?

Formative Feedback Systems and the New Instructional Leadership

These functions help describe how data enter the school, how data are stored, how practitioners use data to set goals and develop plans, what schools put into place as a result of the data, and how students are prepared to generate the next round of achievement data.

We used the DDIS categories to guide our initial data-coding analysis. To make sense of the 300 pages of fieldnotes and artifacts collected across our schools, we used a qualitative data analysis program (NVivo 7.0) to code our data in terms of an analytical scheme based on the DDIS framework. While we came to the data analysis with pre-defined DDIS codes, we soon found that some of the data either would not fit into these initial codes or spanned several of the coding categories. An initial pass through the data using the DDIS coding framework led us to sort our data into the DDIS categories and to identify the key artifacts leaders and teachers used to generate and reflect upon formative feedback information. We then reanalyzed the fieldnotes or narratives to identify the narrative “chunks” that would give us deeper insight into the local formative feedback systems.

Findings

Formative Feedback System Model

This paper reports on the formative feedback processes through which schools generated, reflected upon, and acted upon data related to the instructional programs they had implemented. We found that feedback was given in schools on a number of different levels—from teachers to students, between teachers, and from school leaders to teachers. Much of this feedback was contextual and informal, offered on the occasion of particular incidents or actions. However, our research also revealed the existence of designed patterns of interaction composed of networks of curricula, assessments, and opportunities to reflect and to act, that provided teachers with sophisticated data to guide student learning. Unlike district-imposed formative feedback systems identified by Sharkey and Murnane (2006), the formative feedback systems we found were composed of ordinary aspects of the school instructional program and arranged to fulfill formative feedback functions. Notably, the schools themselves did not identify the systems as distinct from the surrounding instructional practice. Each of the patterns of interaction we observed seemed to involve common kinds of artifacts that generated feedback to facilitate teacher-teacher and teacher-staff interaction about student and program performance. The formative feedback systems we observed in the DDIS schools were all focused on providing more detailed, timely information for teachers on literacy, language arts, special education, or student behavioral programs.

Here, we propose a model to describe the patterns of interaction we observed, then use the case of Pearson Elementary School to illustrate how the model functioned in the context of an elementary school. The three key functions (Figure 1) of the formative feedback systems model are *intervention*, *assessment*, and *actuation*.

Intervention

Interventions describe the intervention artifacts deployed to/for groups of students to improve their learning. Most often, intervention artifacts took the form of curricular materials, such as textbooks, experiments, worksheets, computer programs, etc., that teachers used to

Formative Feedback Systems and the New Instructional Leadership

structure classroom learning. Staff also used an adaptation of the individualized education program (IEP), borrowed from special education, as a tool for customizing instructional and behavioral resources to meet the perceived needs of certain students in and out of the classroom (Halverson & Thomas, 2007). Thus interventions comprise the artifacts school staff use to influence instruction, either curricula materials or IEPs. The learning that results from interaction with interventions is analogous to the *signal* of an classic information processing system.

Assessment

Assessments measure the degree to which students have learned what was intended in the interventions. Formative assessments provide specific information about the degree to which aspects of the intervention succeed or fail, and should to point to how teachers might revise instruction to meet student learning needs. In the formative feedback systems we observed, teachers and staff used either commercially available assessments or designed their own assessments to determine whether the interventions had their intended effects on student learning or conduct (see Figure 1). Assessments are analogous to the *sensor* capacity of an information processing system because they determine the degree to which signals received (estimates of student learning) correspond with the learning goals built into the interventions.

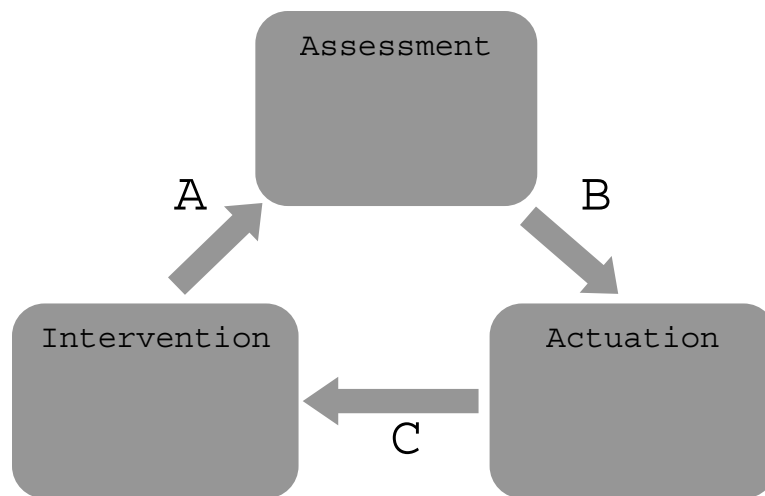


Figure 1. Formative feedback system model.

Actuation

Assessments of interventions, however well designed, merely provide information. In order to turn this information into organizational knowledge, schools need structured spaces to turn information into knowledge. Actuation refers the process through which faculty and staff come to understand, and act upon, the effects of their interventions on student learning. Designing for actuation means setting up legitimate spaces, such as faculty-, grade-, and team-level meetings, for teachers to reflect upon the data and to make decisions about how to alter program delivery for students, or, in cases of significant problems revealed by the assessment,

Formative Feedback Systems and the New Instructional Leadership

how to alter the interventions or the assessments themselves. Actuation spaces combine the signal processing and the controller aspects of the classical feedback systems model. Actuation spaces help practitioners make sense of assessment information, and, with adequate organizational support, provide the opportunity for staff to make appropriate adjustments to the intervention.

Pearson Elementary School Formative Feedback System

The work of instructional leadership is to create structures to support each of the functions of the formative feedback system, and to help teachers and staff link the functions so that formative data can flow through to inform decisions. The leaders and teachers at Pearson Elementary School have organized a tightly coupled, sophisticated formative feedback system around early literacy instruction. Pearson Elementary School, in a rural Midwestern town of 10,000, serves about 300 students in Grades K–6. Pearson was reconstituted 8 years ago in the hopes of creating a better school for children who lived on “the wrong side of the tracks.” Just over 40% of Pearson’s students qualify for the free and reduced-price lunch program. The percentage of students testing at the proficient level on the state exams has grown from 2000 to 2005: in math, from 33% to 92% students proficient and advanced, and in language arts, from 28% to 97%. Under the direction of Principal Eve Meadows, the achievement and the reputation of Pearson have improved so markedly that the school is now faced with an influx of more affluent students and families willing to cross the tracks.

Teachers at Pearson credit the effective use of data for much of their success with improving student achievement. As one teacher put it succinctly: “We use the data ourselves to see student growth.” Another teacher commented that:

What we’re looking at more than just the test, we’re looking at the whole package. You know, what can we all do together as a team to help this child to feel good and be successful in school.

Over the past 8 years, Principal Meadows and her teachers have built structures that allow teachers to reflect on standardized test data and to focus on the particular problems, such as the development of early literacy skills, that thwart student learning gains across the curriculum. In this discussion, we focus on the structures and practices that provided formative feedback to K–2 teachers on their reading and writing programs. In terms of our formative feedback system model, the reading curriculum is the *intervention*, a battery of commercial exams used by Pearson teachers is the *assessment*, and the regular grade-level meetings for teacher reflection and action are the *actuation space*.

Intervention: Guided Reading and Orton-Gillingham Phonics

The Pearson early literacy program is a sophisticated package of curricula stitched together under the guidance of Title I teacher and literacy specialist Charlotte Wagner. The impetus for the Pearson literacy program was a district (and state) press to improve the quality of K–2 reading instruction for all students. The new initiative, adopted in 2000, was in stark contrast to, as Wagner put it, “the early years when we looked at what the child couldn’t do.”

Formative Feedback Systems and the New Instructional Leadership

Pearson’s literacy program supplements a Guided Reading foundation with Orton-Gillingham phonics instruction. Wagner stated that the Fountas and Pinnell (1996) Guided Reading book “is our Bible.” Wagner was trained as a Reading Recovery (RR) teacher, but found RR too expensive and intensive to serve the needs of all students at Pearson. RR was also provided by pulling students out of their classrooms, which meant that “a child could have a classroom teacher, a different Title teacher, and another RR teacher.” She found the Guided Reading program to be “an absolutely excellent” alternative to RR for teaching reading, and helped the K–2 teachers adapt some of the RR writing materials into the literacy program. Guided Reading teachers help small groups of students use the contextual and visual cues in a book to understand the meaning of stories. Guided Reading students progress through a series of texts organized according to demonstrated reading levels. Student grouping is determined by a series of quick assessments, or *running records*, that each teacher is expected to conduct to track student progress.

Guided Reading, though, proved difficult to implement with new readers who struggled with simple phonics skills. All Pearson reading teachers have attended workshops on the Orton-Gillingham approach to phonics teaching. Orton-Gillingham involves daily practice in sounds and word-decoding skills to prepare students for book reading. The Guided Reading sessions could then be used as diagnostic sessions to identify the kinds of phonics skills students were missing. The common staff training and commitment to Guided Reading and Orton-Gillingham helped provide program focus for students so that, in Wagner’s view, “now the language is common, and even our struggling readers understand what we are talking about.” (See Figure 2.)

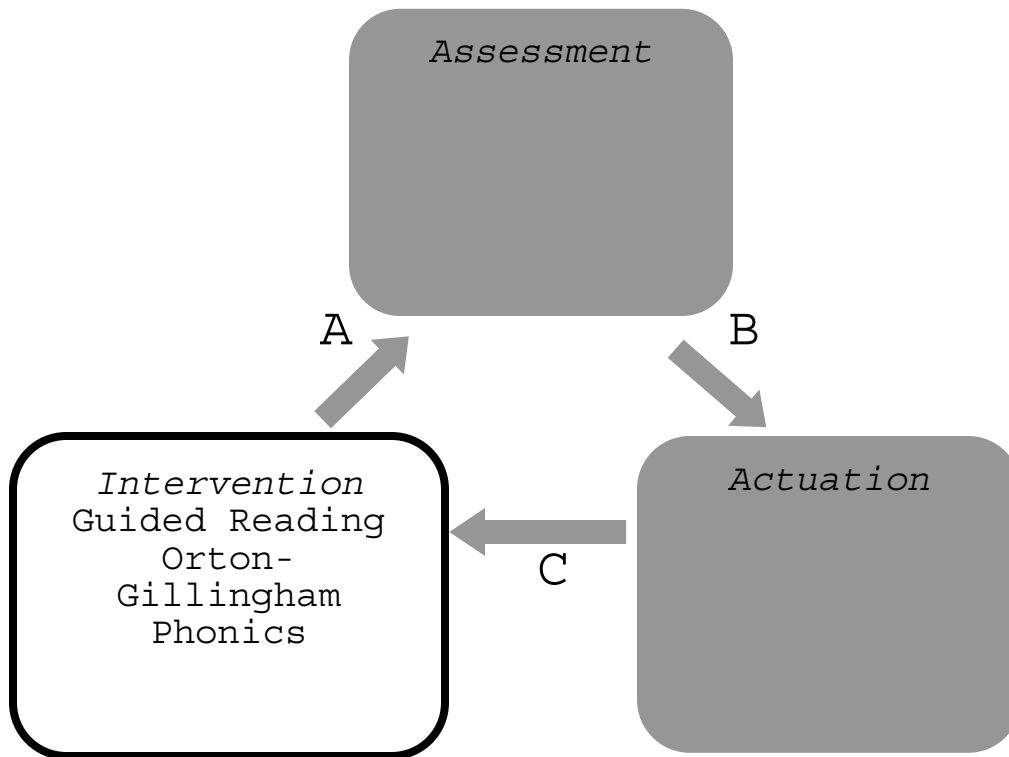


Figure 2. Pearson interventions.

Formative Feedback Systems and the New Instructional Leadership

Principal Meadows played a key role in setting up the conditions necessary for the new literacy program to succeed. Wagner was encouraged to redesign her role in the school from providing services to Title I students through a pullout program to serving students in the context of regular classrooms through team teaching with the Grade 1–2 teachers. Meadows also pushed for providing access to the training necessary for all teachers to teach all students. In Meadows' view, the curriculum is important, but the teachers are the critical resource:

The thing we can't forget is that I also have a highly trained workforce in the area of reading, and I've made that clear to the board that let's not forget the person that's delivering the instruction. Many of my teachers have gone back and gotten master's degrees in reading, because they realized that boy, that's just everything.

She explained that “the outcomes should be that all kids can read, and we know that they're going to all get to that point in a different way. Lots of different strategies are necessary.” Meadows provided support for Wagner and the Pearson teachers to figure out which strategies worked best for different students.

Assessment: Running Records, DRA, Buckets, and Notebooks

Assessing the effects of the literacy program on student learning is a critical component of Pearson's literacy program. As we might expect, Wagner and the Pearson teachers brought together several kinds of commercial assessments in order to better measure the learning needs of students. We were surprised, however, to find a range of low-tech data collection artifacts that served as data-recording tools, including space for anecdotal student information. The team developed a decidedly low-tech process for recording, storing, and distributing the assessment data in plastic classroom buckets and student assessment binders. (See Figure 3.) Principal Meadows commented that, with this assessment system, “now we actually know how well they're [students] doing or what their areas of weakness would be in.” The assessment tools provide benchmarks for student learning and a way to identify the specific areas in which students require instruction. Here, we highlight the main assessment tools in the Pearson literacy formative feedback system: running records, developmental reading assessments (DRA), buckets, and assessment binders.

Running records. Running records (Clay, 1985, 2000) play a central role in the Pearson literacy formative feedback system. As Charlotte Wagner explained: “Running records drive our instruction, letting us know where we need to go next.” Teachers use these widely adopted assessments to test a student's developmental reading level by recording the errors, self-corrections, pauses, and questions as the student reads aloud from a book. The teacher selects a book that s/he perceives to be at the student's reading level, and as the child reads a passage, the teacher annotates another copy of the text marking incorrectly pronounced words. The teacher also attends to how the student is using meaning, structural, and visual cues to read the passage. After the reading, the teacher checks for comprehension. The student retells the story and is questioned about the characters, main ideas, supporting details, sequence of events, setting, plot, problem and solution, and response to text-specific vocabulary and language.

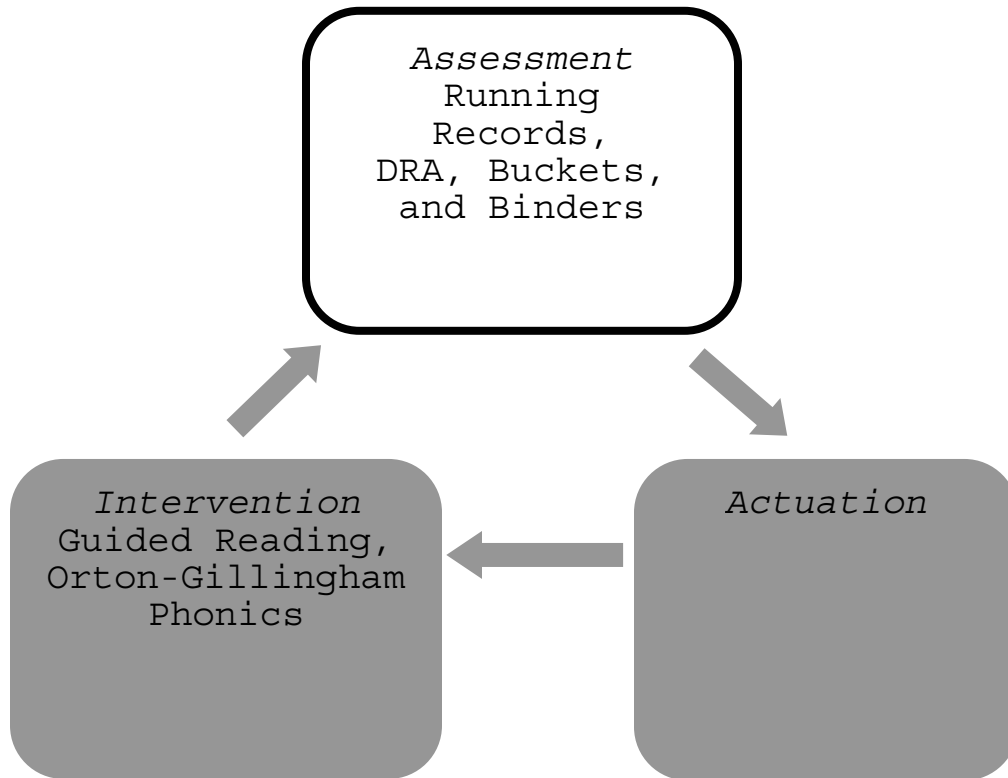


Figure 3. Pearson assessments.

Data from running records plays a central role in guiding differentiated student instruction. Wagner feels that the running records process is directly linked to the Guided Reading curriculum—the tests measure the kinds of skills the teachers teach, and the tests provide the kinds of information teachers can use to revamp their teaching and to reorganize student groups. Each running records administration at Pearson takes from 20 to 30 minutes and the results are compiled in the student data notebook. In order to relieve teachers of the time required to administer the assessments, and also to ensure the reliability of the assessment process, Wagner assesses each child in the K–3 program weekly. These assessments provide a profile of student reading progress over the course of the year.

Developmental reading assessments (DRA). The Pearson school district mandates the developmental reading assessment (DRA) to track student accuracy of oral reading and comprehension. Like running records, the DRA involves reading a series of books with children to gauge accuracy, fluency rate, and phrasing. Students then retell the story in the book to teachers to measure comprehension. The tests are given three times per year, and the scores are reported to the district. At Pearson school, the DRAs have had a mixed reception. Because the data goes to the district level, some staff feel that the DRAs are used to evaluate teaching quality rather than to provide formative program information. However, Wagner has repurposed the DRA scores to triangulate the running records and the state reading exam. As she puts it: “the running records help determine where kids should be on the DRA, and the DRAs predict the

Formative Feedback Systems and the New Instructional Leadership

(state exams). We are rarely surprised.” As with the running records, Wagner conducts all the DRA assessments to ensure reliability.

Buckets. Plastic buckets hold all the materials teachers need to work with each group of students. The humble buckets hold the notebooks belonging to students in a particular reading group, lesson plans, past and present group affiliations for the student, a copy of the book (or phonics exercises) the student is using, and daily anecdotal teacher observations. The bucket plays a key role in coordinating assessment and instruction. Books are added and removed as students reach new reading levels, and all of the material for a student is moved to another bucket if the student shifts groups. Wagner and the teachers can coordinate their perspectives on the instructional program for each student by reviewing the contents of the buckets for a quick, tangible check on their agreement about student grouping decisions. Special educators and other specialists can quickly assess the progress of a particular student or group by reviewing bucket contents.

The student notebooks track student progress through each lesson and follow the student throughout their time at Pearson. Examining the history of group assignment provides a quick longitudinal profile of student progress. This function of the notebooks helps the teachers make the critical connection between assessment and instruction: to determine which aspects of which lessons contributed to (or hindered) reading progress as measured by the assessments. The notebooks also serve as a place for teachers to make anecdotal observations of student learning and behavior. The buckets provide a low-tech but valuable repository for classroom formative assessment information. One teacher described how Wagner uses the notebooks to help with “the kids that aren’t getting it” by doing

daily reports or agendas with all of those kids. . . . And she checks out with them every night. If she weren’t doing that, that would be me doing that in addition to the—to regular ed students that are on—that need that kind of extra support too.

Wagner reviews the bucket contents weekly to provide essential teacher support by helping to creating specific learning suggestions for individual students, ranging from ideas of assessments, to instructional strategies, to regrouping ideas. The buckets serve as a common frame of reference for the teachers and Wagner to understand and address student learning needs.

Assessment binders. A collection of three-ring plastic binders, one for each classroom teacher from each of the past 8 years, grace the top of Wagner’s office file cabinets and bookshelves. These assessment binders provide the storehouse of instructional data to guide the reading program. Each week Charlotte Wagner transfers the information from the buckets, running records, and DRA into these low-tech data storage tools. Wagner has developed several kinds of forms to report on the vast amount of information stored in the binders. The forms include:

- Records of individual student performance tracked over time;
- Collections of anecdotal information on students, such as parent conferences, illnesses, special education referrals, and behavior issues;

Formative Feedback Systems and the New Instructional Leadership

- Forms to track the progress of individual students over the year that end up in the student record file;
- Forms that track the progress of each student in a given class, in terms of words learned, sentence dictation, and phonics.

Although the binders are available to teachers, Wagner is responsible for entering the data and for generating the reports. The assessment binders supplement teachers' perspectives on student learning and give the teacher community a representation of which skills students have and have not mastered. Teachers can track whether their insights and interventions with particular students, or with groups, influenced learning over time. As Wagner said, the assessment binder provides a means to "look at what the child can do and how they have grown."

The binders are used to address programmatic and staff development issues as well. Principal Meadows uses the binders to track how teachers are progressing with students. For example, she uses the binders to track whether the interventions put in place for particular students have worked over time. Several years ago, the Pearson faculty integrated Guided Reading into their summer school program, where struggling students are regularly assigned to catch up with their peers:

So, what I do, is I take . . . this year's teacher's paperwork, and I look at last year in May at the results of their testing. And one thing I saw, for example, Gale [pseudonym], when we left in May, he was at a Level 12 in the developmental reading assessment, and in September he was at a 16, and so my first question was, did he go to summer school and take Guided Reading? And the answer was yes. And so I went through and [found] a couple of them on the opposite end, okay, hmm, I see that this child was a Level 4 when we left in May, which is low, and that's certainly a concern, and he's a Level 3 now. Did he go do summer school? No, the parents refused to take him.

The binders furnish a longitudinal record for Meadows to use to focus school resources on struggling students. Once identified, Pearson staff can develop appropriate action plans:

That begins to draw a picture. Number one, let's look at some concerns in the backslides and the increases, but also we can begin to have that dialogue with parents about the difference summer school can make, particularly the intentionality of Guided Reading in summer school for your child, so they don't have the backsliding. And so taking the information from last year, laying this year's information on top of it, and beginning to look at kids and talking individually about where they're at.

Meadows also uses the binders to structure her end-of-the-year conferences with teachers. They consider the patterns of student learning and discuss new plans for some students. Since the binders allow K-2 teachers a common reference for working together on student literacy, plans made for one year can be followed up in the next. Finally, the binders also serve to mediate parent communication. In the past, teachers found it difficult to show parents how the Pearson literacy program worked. Sharing the binder records with parents helped confirm her hunches about whether parents or teachers were giving students proper support. Wagner stated, "I now have the data to show the growth of that child."

Actuation: Creating Spaces for Deliberation and Action

Pearson’s leaders have created multiple legitimate spaces for teachers and staff to reflect upon the assessments and decide whether to alter subsequent interventions at the student level or to make broader curriculum-level changes. These actuation spaces (see Figure 4) afforded designated opportunities for teachers and staff to adjust instruction to better meet the needs of individual students, or, more infrequently, to make adjustments to the interventions themselves through adding new program components. The main actuation spaces we observed included (a) Wagner’s team-teaching schedule, (b) grade-level faculty meetings to discuss data and student placement, (c) monthly staff meetings, and (d) Implementation days.

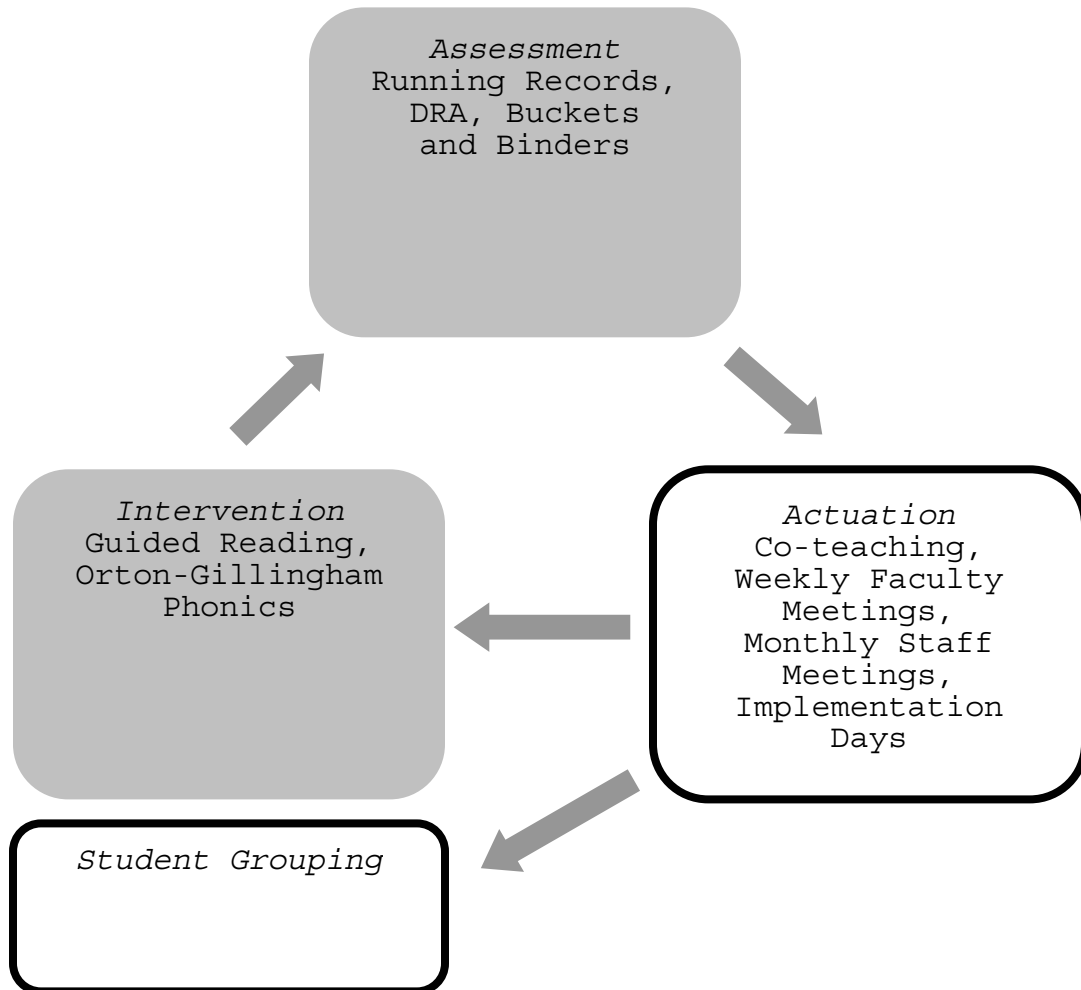


Figure 4. Pearson actuation spaces.

Wagner’s team-teaching positional role. People, of course, cannot be considered artifacts from a distributed leadership perspective, but positional roles provide malleable artifacts for leaders to use when shaping school instructional practices. At Pearson, Meadows and Wagner have redesigned Wagner’s Title I teaching role to maximize her opportunities to interact directly with teachers and students. Wagner’s reconfigured position was designed to spend ½ day in each teacher’s classroom during the reading lessons. This allows her to spend one day every 2 weeks

Formative Feedback Systems and the New Instructional Leadership

in each teacher's room. Wagner and the K–2 teachers have worked to arrange classroom instruction to allow for collaboration. During the reading lessons, students are broken into three small groups of four or five students.¹ The classroom teacher does reading lessons with one group, Wagner does reading and assessment with another group, and the third group works in Learning Centers on enrichment activities such as computer-based reading. The groups rotate after 20–25 minutes. This allows Wagner to work with each student, to model instruction, and to see how the teacher engages in interaction with student groups. Having Wagner in the classroom creates an authentic team-teaching environment. As one Pearson teacher noted:

I have another person in my classroom whom I respect, who is sharing information with me, who talks to me on a regular basis back and forth, on what we see in the children. Whether it's from testing, whether it's from them in a group. I'm sharing that with someone.

Here, we provide an example to illustrate Wagner's classroom interaction with the teacher and with students. Language arts instruction at Pearson takes place in the morning. Four students gather at Wagner's table and are each given a copy of the same book from their buckets. Wagner asks the students to examine the cover in order to predict the plot. She directs students to do a "picture walk"—to look at the pictures in the book, reporting what they see. She records each child's response on a small whiteboard: "dinosaurs," "important stuff," "a knight." She seizes this opportunity for students to notice the contrast of *knight* with *night*. "What is tricky about knight?" she asks. A girl replies, "The silent K." Wagner encourages the students to stay focused on the picture prediction exercise to "get your minds set—no reading yet Sonia. . . ." Wagner asks the students to turn to page 31 where she highlights a "tricky" word—*instead*. To help students sound out the word, she asks the children, "Do you know another word like that?" The students appear to be struggling with the task, so Wagner reminds them to "think about what they (already) know." Returning to the storyline prediction task, Wagner asks, "What would you do if you saw a dinosaur?" One boy says, "run away, especially if it was a T-Rex!" Wagner concludes the picture walk activity by informing the children about their upcoming activities on Tuesday (running records) and Wednesday (hot seat retelling). The students begin writing "I predict stories" based on what they think will happen in the book they looked at with Wagner.

Meanwhile, the classroom teacher works on phonics with a lower level ability group of three students. The teacher greets the group—"Did you have good weekends?"—and begins flipping through phonics flash cards. As she flashes the cards, students trace the letters with their fingers, write them three times, and say the sound. The teacher immediately corrects student errors as they go, and the children repeat their corrected responses before moving to the next card. The other students not in one of the adult-orchestrated groups work on specific tasks differentiated to meet their individual needs, such as computer-based reading, story prediction and writing, and kinesthetic flashcards with feedback.

In addition to her instruction work, Wagner makes other classroom visits to administer running records and DRAs. Administering the assessments provides Wagner a firsthand view of which aspects of the reading program and instruction require more focus. For example, the running records will record if a student is actually reading the words they are seeing, and

¹ Pearson participates in Student Achievement Guarantee in Education (SAGE), a program designed to limit class sizes to no more than 15:1 in Grades K–3. SAGE schools contract for state aid equal to \$2,000 for each low-income child in the grades served by the program.

Formative Feedback Systems and the New Instructional Leadership

Wagner's observation of the student will provide contextual information to help her interpret the test results. The teachers appreciate the role Wagner plays in the assessment process. As one teacher remarked:

It's one person giving the same test to all the children in the same way. You know, if we each gave the tests as a classroom teacher, probably all three of us would give it in a different way. It's more valid, we feel, if one person is giving it the same way to all the children.

Wagner spends 15–20 minutes per student administering running records on a weekly basis for each first grader, and every other week for second graders. At the end of her assessment sessions, she speaks briefly with teachers to note surprising results and to keep teachers updated.

Faculty meetings. The schedules of the K–2 faculty members are organized to provide two opportunities to discuss the literacy program with Wagner. Wagner meets with several of the individual teachers each Friday and meets monthly with all the teachers to discuss student data and progress. A floating sub relieves teachers to meet with Wagner. The teachers discuss the data on children and the reading program, and they often add anecdotal observations, about personal events such as divorce or absences, that may be influencing the child's learning.

The individual teacher meetings give Wagner and the classroom teacher an opportunity to compare notes about individual students and to discuss whether to change instruction or group membership for students. Wagner has the opportunity to meet with each teacher for an in-depth discussion every 3 weeks. One teacher explained how she made sense of the assessment data in these meetings:

You're assessing them [running records], which is all [done] individually . . . you read those running records and from that then you decide how you're going to give your instruction the next day. And then you might change the groups because they might, this kid got it, but this kid didn't.

Though Pearson has worked to move away from ability grouping in the primary grades, Wagner stated that the school is not ready to implement differentiated group instruction.

[According to Guided Reading], groups should be flexible to teach the specific skill, bringing together high- and low-achieving kids to work together on specific skills. We are not there yet . . . we look at what strengths they have in common . . . but for the 15 minutes a day [of reading instruction], they are somewhat ability-grouped.

The constant adjustment of group membership, added Wagner, blunts the negative effects of ability grouping and leads to student shifting in and out of groups based on achievement. The weekly meetings with individual teachers give Wagner and the teachers a better perspective on what students know. This gives "kids a chance to feel comfortable with what they are doing. They are able to say 'I am a reader.'" The daily and weekly data collection provides the foundation for the monthly conferences she holds with teachers.

Monthly staff meetings. Each month, Wagner organizes an hour-long meeting to review the literacy program with all the teachers, the special education staff, guidance counselors, and the instructional aides. These meetings give staff a chance to review what they know about the

Formative Feedback Systems and the New Instructional Leadership

learning problems of individual students. One teacher commented on how these meetings allow teachers to

talk with . . . other people in our building, who maybe don't work with the kids every day in the classroom. That is, an outside voice. The guidance counselor, who doesn't see the kid every day, might have a different interpretation about things than we do. . . . [She] may know that other things are going on, that may help explain things that have been done in the past, [the] different testing that may have been done. Or that they've talked about doing different testing.

Often, the monthly meetings take the form of study groups that focus on the issues involved with Guided Reading, phonics, or the writing program. During study groups, teachers read and discuss a research article relevant to the program area and discuss what they find to be problems with the current program. These meetings also allow the teachers to compare notes on children in their classes and provide an occasion for the lower grade teachers to relate comments about students who continue to struggle in the higher grades. For Wagner, the study groups also help the staff keep the original design of the reading program in mind.

That is one of the reasons why we have the study groups. Our personalities can overtake the program. [Study groups] remind us of those essential principles of the program [that] still need to be met.

The strong personalities on the Pearson staff tend to individualize the instructional program, and the study groups preserve a common instructional foundation agreement among staff.

Meeting together as a group helps Wagner detect programmatic issues. For example, after putting the teaching and meeting system in place for the first several years, Wagner noticed the effects of the directions teachers were giving for instructional activities.

We found that one classroom teacher was saying 'sound it out,' another teacher was saying 'stretch it out,' I was saying 'say it slowly.' Our top kids were figuring it out, but the kids who were struggling were saying 'I gotta do something else?' They didn't get it.

Participating in the classrooms and the faculty meeting actuation spaces allowed Wagner to see the cumulative effects of these subtle differences. It was not the instruction that was at issue as much as student confusion about the directions. A common focus on Guided Reading has established a common instructional language so that "the struggling readers could figure out that the skills sought by different teachers were the same."

The whole group meetings also help identify instructional design gaps in the program. The phonics program implementation provides an excellent example. Several years ago, Wagner and her teachers were finding the middle- and high-achieving students were responding well to the Guided Reading program. The low-achieving students, however, continued to struggle with basic phonemic skills. Several of the teachers had noticed this program deficiency, and one special education teacher recalled an Orton-Gillingham phonics training workshop she attended many years before. Wagner and the literacy teachers liked the concrete, easy-to-implement Orton-Gillingham phonics program, and agreed to try the program with students. Rather than allow the teachers who had already received the training to treat students through a pullout phonics program, Wagner worked to secure the resources for all teachers to attend the workshops

Formative Feedback Systems and the New Instructional Leadership

so that each staff member could provide consistent phonics instruction in the course of regular literacy classroom teaching. This example shows that the actuation spaces in the Pearson literacy formative feedback system do more than help teachers customize learning for students; they also provide the space for program review and alteration for all students and staff.

Actuation is part of an iterative improvement process. Actuation closes the system by translating interpreted assessment data into student-based interventions. Charlotte Wagner's position serves as the engine for actuation by compiling and knowing individual student data. Her ongoing work with teachers and intensive monthly meetings provide a means of differentiating instruction to meet the specific needs of each student. Pearson's reading program uses formative feedback to adapt its programs to the needs of the students.

Implementation days. The final designed actuation space that relates to the Pearson early literacy formative assessment system is a series of school-wide *implementation days*. These quarterly half-day in-services help teachers and leaders connect local instructional efforts to the school-wide goals that are determined through annual review of the standardized testing data. Although not specifically targeted toward reading goals, the implementation days allow the reading teachers to understand how their instructional efforts map onto standardized tests, thus creating a bridge between the internal formative feedback system and the summative data system for which the school staff was held accountable.

The implementation days provide an actuation space to consider how well current instructional practices are assessed on the state test. For example, during the final implementation day in the spring of 2005, teachers noticed questions that addressed image and sentence fluency skills not addressed by the instructional program. During the meeting, each teacher received a packet of data tables and charts summarizing the standardized test scores from the fall. Ninety-five percent of all students tested at proficient or advanced levels, including all of the special education students. Principal Meadows opened the discussion by announcing, "As we can see from our scores, [we] did great again this year, and I think we can trace it all back to our program and staff." After reviewing some of the general findings, Meadows led teachers into an examination of the student performance on specific types of questions. She asked a fourth-grade teacher to "take us into the test." Some of the questions asked students to combine sentences, and one teacher commented that "our daily oral language exercise doesn't support that. How can we put that in?" The fourth-grade teacher noted that there were several resource books available that focused on issues like sentence fluency. Meadows and the teachers kept coming back to the skills that teachers did not previously address in the literacy program. A special education teacher mentioned that "our fourth graders are doing well, but that's not to say they always will." Charlotte Wagner added that "some of the [first- and second-grade teachers] do sentence writing, but the sentence part of the literacy is weak." Another teacher added, "[fourth-grade teachers] do this in our own writing instruction, but it doesn't look like this." Meadows summarized the discussion with the comment that "students might have the knowledge, but if the presentation is foreign, there might be a way to remove the roadblocks."

This discussion was one of several that took place during the meeting that showed how the Pearson community used the state tests to illuminate gaps in the instructional program. Starting with a strong, shared understanding of the literacy program helped teachers recognize

Formative Feedback Systems and the New Instructional Leadership

what was missing. Implementation days allow teachers to keep tabs as students progress through the grades. One first-grade teacher mentioned how the implementation days allow teachers to

break the tests apart, child by child. We go and see . . . who is struggling, who we need to red flag—we all need to watch this child. My opinion is asked, even though this is a second grader and they aren't with me anymore. They still look to me to say anything I might know from this child. So, I mean it's respect . . . we watch these children all the way from kindergarten up to sixth grade. And we talk about these same children all the way through. We compare their data from where they started to where they ended up.

Implementation days are actuation spaces that allow teachers and staff to use assessment data to draw together their experiences and insights in order to develop the best learning opportunities for their students as they progress through the grades at Pearson.

Summary

Pearson's formative assessment process works, as confirmed by the data which shows the students advancing through the expected levels. Wagner said, "For NCLB, we do have AYP data, they may not reach the benchmark, but they are growing." While the school has not succeeded in helping every student meet benchmarks, each student is showing growth. Using data to monitor classroom-level interventions allows a Pearson staff member to keep constant tabs on student reading learning and provide tasks that fuel student progress. The professional impact of the formative feedback process is summarized by one of the teachers: "I think that the data use has made us better teachers. I think it's made us more aware of individual students . . . on where they excel and where they need help." Being a better teacher at Pearson means using data from individual students to identify student strengths and weaknesses. Since the goal is that each child learns to read, the professional expectation is that instruction responds to the student. This results in teachers' learning how to help every student learn to read.

While the program is successful in teaching students to read as measured by state standardized tests, ongoing attention to how reading is taught to specific students constitutes the heart of the school's formative feedback practices. This attention to the process of formative assessment and its results are described by Wagner:

Right now, we are showing growth. We fine tune it [the teaching of reading] every year . . . and now more and more children are meeting the benchmarks. . . . Last year we had two children in third grade who were basic, they missed proficient by one point.

None of the students at Pearson tested below the proficient level last year.

Discussion

This case of a formative feedback system illustrates how Pearson's leaders and teachers intentionally designed interventions, assessments, and actuation spaces to provide customized information loops for informing teachers and staff about student learning. At Pearson, we observed how teachers and staff worked with the tight linkage between literacy program interventions, assessments, and actuation spaces. The scope of the intervention-assessment linkage figured prominently in the design of formative feedback system. We observed similar

Formative Feedback Systems and the New Instructional Leadership

exclusive foci on literacy instruction in other DDIS schools. We did not observe similar connections in any other subject area in the schools.

This focus on the literacy instructional domain has several possible causes. First, each of the schools we studied discussed the difficulties involved in teaching children to read. Reading was a primary emphasis on the state exam, and it was difficult for many of the schools to meet accountability standards without the ability to monitor and adapt reading instructional strategies. Math curricular programs, on the other hand, typically provided a rich set of formative assessments in the textbooks. Once teachers were trained in the math program, they could teach students at a level with sufficient competence to help the school avoid accountability consequences. Second, it is expensive, in both personnel cost and time, to maintain a formative feedback system. In Pearson's case, faculty interaction in the early grades was dominated by literacy-based actuation spaces, and Charlotte Wagner's central role was to collect data, engage in instruction, and organize and conduct interaction about early literacy in actuation spaces among the six first- and second-grade faculty members. Her position represented a significant percentage of the total instructional support resources available to the Pearson staff, and the choice to commit her service to early literacy shifted resources away from other school priorities. Finally, the focus on literacy may have reflected the staff capacity to engage deeply in instructional issues. The Pearson staff, for example, had six teachers with advanced degrees in reading, but did not have anyone with advanced training in math or math education. While it is true that four members of the Pearson staff obtained advanced training in reading after the school made a commitment to early literacy instruction, it remains an open question whether similar formative feedback systems could be built in the school without the kinds of deep subject matter knowledge that facilitates sophisticated learning diagnoses and teaching interventions.

Here, we discuss several interactive features of the formative feedback system model that bring out conditions of system operation as a whole. (See Figure 5.) The functions of the formative feedback system describe the main activities in which teachers and leaders engage to create meaningful classroom information. The system linkages seem to be critical to allow the kinds of information that would help inform teaching to flow through the system. Here, we highlight the issues involved with the three system function linkages: *intervention-assessment*; *assessment-actuation*, and *actuation-intervention*.

Intervention-Assessment

The key design question related to the intervention-assessment link is: Do teachers have faith that the assessment measures the intervention? Aligning interventions and assessments creates a coherent signal that staff can hear and act upon. The link between intervention and assessment at Pearson, for example, allowed teachers to see how their daily instructional practices could be measured by shared assessments, and in turn, how the assessment data could be used to make instructional decisions. Porter (1995) argues that unless there is a tight match between what is assessed and what is taught, the assessment results are meaningless, and the resulting decisions are potentially harmful. If the measures of learning do not follow directly from instructional practices, teachers may have difficulty determining how to interpret the resulting signals in terms of teaching practices.

Formative Feedback Systems and the New Instructional Leadership

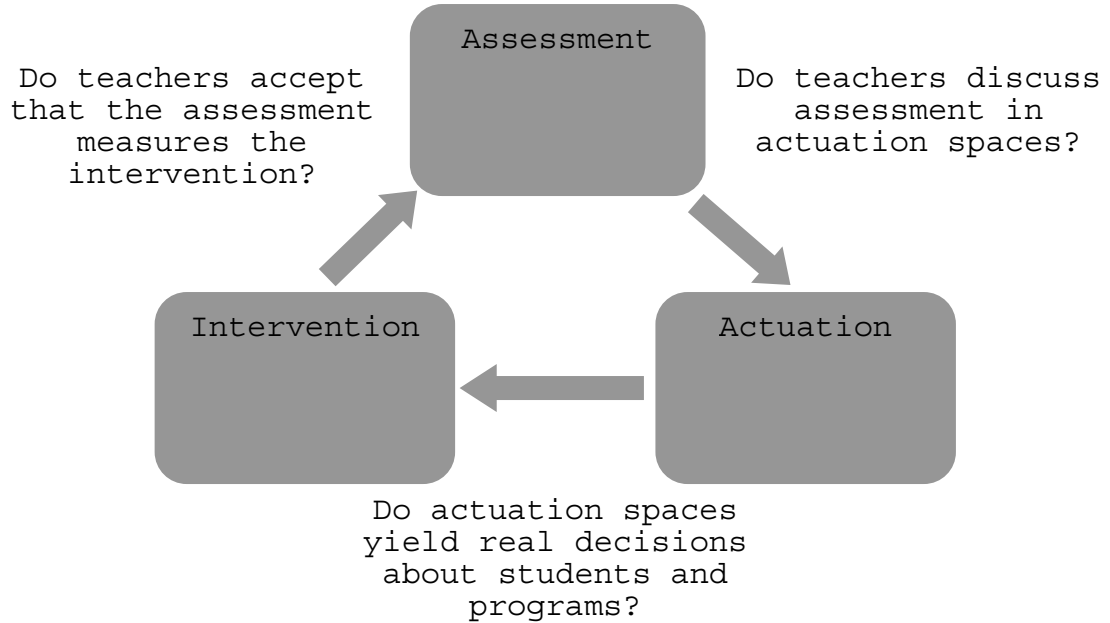


Figure 5. Formative feedback system linkages.

This provides a chicken-or-egg design question: Which comes first, the assessment or the intervention? Starting either with the test or with the curriculum requires schools to construct the link to the other. In one of the DDIS schools, for example, the school started with a commitment to a common test. In this school, the district decided to invest in the Northwest Evaluation Association's (<http://www.nwea.org/>) Measures of Academic Progress (MAP) program. MAP is a computer-adaptive test developed to provide timely feedback on how well students are learning in terms of state standards. Since the test is linked to state standards and not to individual curricula, the actuation spaces of this school in our study became dedicated to constructing the links between the MAP test and the everyday curriculum.

In another DDIS school with a long tradition of teacher classroom autonomy, teachers and school leaders had worked to develop a common approach to instruction based on the Four Blocks Literacy Model (Cunningham & Allington, 1999). The Four Blocks approach allowed teachers to customize their own assessments to determine student progress. Developing a formative feedback system, however, meant that teachers would now need to commit to a common set of assessments in order to monitor student progress through the academic program. In the absence of a coherent link between the intervention and the assessment, the staff interaction in the actuation space focused on determining how the tests fit the curriculum or how the curriculum fits the tests, rather than how the information might be used to make decisions about student learning.

The alignment of assessments and interventions at Pearson was the result of a long process of program design through the development of a strong professional community empowered to make instructional decisions. Both Charlotte Wagner and Eve Meadows related that developing a common reading intervention developed slowly over the prior 10 years. As

Formative Feedback Systems and the New Instructional Leadership

Wagner described, in the early years at Pearson teachers categorized students into “bluebirds and buzzards”—that is, children who could read and children who couldn’t. The buzzards were identified as special needs students and pulled out of classrooms for supplemental Reading Recovery lessons. The old pullout service delivery model effectively ensured that classroom teachers could continue to teach as they had always taught and could resist collaborative curriculum design. Shifting priorities to address the needs of all students required teachers to accept a new, collaborative approach in designing the literacy teaching program, Guided Reading. This collaborative approach allowed teachers to share their expertise and agree on the best approach for teaching all children. The link between intervention and assessment emphasizes the importance of a common vocabulary for instruction in a school. This slowly emergent common agreement about curriculum allowed teachers to discuss how their classrooms were progressing, and established the conditions for seeking out common assessments to determine whether the instruction program was working.

Assessment-Actuation

Do teachers actually discuss assessment data in actuation spaces? Taken in the ordinary sense, an *actuation space* is a technical term for the homely faculty or a grade-level meeting. Researchers and practitioners have long noted the necessary, but insufficient, value of creating spaces for faculty discussion (Peterson, McCarthy, & Elmore, 1995). At Pearson, the available faculty interaction space was radically redesigned for teachers to consider assessment data. When aligned with interventions, assessments can act as *sensors* to detect students learning. The alignment between assessment and actuation determines the *signal processing* capacity of the formative feedback system to process the signal in terms of student learning. Unlike a mechanical system, of course, a social formative feedback system requires appropriate interaction to guide practitioners in making sense of the signal. This sense-making capacity of the formative feedback system can provide the data and a context for understanding the meaning of the assessment data. Research on sense-making in schools often emphasizes the contrast between the policy-signal and the local context for understanding (Spillane, Reiser, & Reimer, 2002; Weick, 1995). Creating a strong link between assessments and actuation spaces is one way that school leaders can influence practitioner sense-making so that teachers are able to act effectively on local practice in terms of the local policy context (Coburn, 2005).

The Pearson actuation spaces helped teachers and leaders make sense of the assessment data at multiple levels. Implementation days helped teachers to understand how the outcomes of instruction were reflected in the standardized test scores. The team faculty meetings provided an overall picture of how students were progressing through the instructional system; the individual meetings allowed Wagner and the teachers to trace progress and to make instructional decisions about individual students within the constraints of the classroom. Finally, Wagner’s team-teaching efforts gave her intimate access to how students reacted to instruction and performed on the formative assessments. The time Wagner was willing to commit to classroom instructional efforts helped legitimize her role with the teachers and helped teachers to accept her efforts as a valuable contribution to the instructional process. These multilevel social interactions created redundant opportunities for teachers to use the data to question instructional practices and to create the types of professional obligations that characterize professional communities (Halverson, 2003). Such professional communities can have a strong influence in shaping how

Formative Feedback Systems and the New Instructional Leadership

information is understood and acted upon in schools (Louis & Marks, 1998; Kruse & Louis, 2006).

Actuation-Intervention

The link between actuation spaces and intervention describes the *controller* function in the formative feedback system. The controller function allows the system to act appropriately on the feedback information. In the case of a school formative feedback system, the controller function enables practitioners to adjust the instructional program. In the Pearson case, we found that the controller function of the actuation spaces allowed for another kind of decision as well—student regrouping. The regrouping decisions represented micro-adjustments to the intervention, allowing teachers to fit their instructional strategies and materials to the needs of individual students. Efforts such as adding the Orton-Gillingham phonics program to adjust the intervention were embraced by the Pearson staff as a result of the controller function between actuation and intervention.

School leaders play a central role in enabling the controller function of actuation spaces. Since teachers are typically focused on classroom instruction, it falls to formal and informal instructional leaders to establish the conditions for teaching and learning (Spillane et al., 2004). We have seen how, in the Pearson case, part of establishing these conditions was allocating financial resources for staff training and time resources for opportunities to meet. Meadows and Wagner designed the formative feedback system to provide teachers with the kinds of information necessary to make decisions about literacy instruction. Empowering teachers and staff to make substantive decisions about the intervention was an important aspect of leadership practice. As we have seen with Thorn's (2001) discussion of administrative and instructional information systems, school leaders often frame information collection activities in terms of summative assessments of system functions rather than formative assessments of system processes. Leaders able to construct formative feedback systems need to understand the difference between summative and formative feedback, and need to be able to maintain separate systems that can yield relevant information on both goals. This two-tiered approach to information design allows instructional leaders to hold teachers accountable for the performance of the school instructional system while, at the same time, designing formative feedback systems and empowering teachers to make the kinds of decisions necessary to improve classroom teaching and learning.

Conclusion

At Pearson Elementary, the staff and administration came to an agreement on the shared curriculum and shared assessment as a way of linking assessment and instruction. Smith and O'Day (1991) have argued that curricular narrowing is a condition for aligning standards, curriculum, and assessments so that children can be tested on what they learn. This design trade-off between control and autonomy is one of the most important consequences of a high-stakes, standards-based accountability policy. Many schools at risk of facing accountability consequences will opt to develop internal systems that provide more control over the instructional process. The Pearson case illustrates how a large part of this control comes through the function of a formative feedback system. Pearson's teachers and staff needed access to

Formative Feedback Systems and the New Instructional Leadership

common assessments to inform specific decisions about student learning, which, in turn, narrowed the curricular options available in the teacher's instructional repertoire. The strength of the formative feedback system is shown by Charlotte Wagner's comment that she is "rarely surprised" when the state test score reports come; but the trade-off for this predictive ability is a narrowed curriculum.

This discussion of the role of formative feedback systems in schools is clearly limited in scope and in effect. Due to the case study approach we adopted, we were not able to pin down the specific effects of formative feedback systems on student achievement. We were not able to directly compare schools similar in other ways, save for their formative feedback systems. However, we were able to identify where formative feedback systems cropped up in our schools, and to compare the ways in which the discourse changed in those domains in the schools that had formative feedback systems and in those that did not. We found formative feedback systems considerably constrained the range of information about instruction discussed by faculty. We hypothesize that there may be an inverse relation between scope of instructional information available to faculty and the quality of data-driven discussions: the more data are available, the more practitioners feel the need to debate the quality/meaning of the feedback, and the less effective the feedback system becomes. These and other similar hypotheses would need other methods better suited to careful comparison of disparate research sites for adequate confirmation.

We are not claiming the components of formative feedback systems are particularly innovative. Thousands of schools use phonics, Guided Reading, running records, faculty meetings, and three-ring binders to conduct their business. We are also not arguing that the leadership or the organization of resources at Pearson is somehow unique. Our argument is designed to make a more modest, realizable claim. From these humble components, school leaders and teachers are constructing powerful formative feedback systems that provide sophisticated, carefully selected information that enables schools to customize their instructional programs on the fly. Describing a model that captures these designed feedback loops of instruction, assessment, and actuation can help researchers document these practices so that others can understand how to organize feedback systems in their schools. In the 1990s, businesses and schools around the world went in search of the elusive learning organization. In the 2000s, we find local leaders and teachers constructing the building blocks of genuine learning organizations in early literacy programs. Perhaps the key to making these practices accessible to all schools begins with the simple step of providing a common vocabulary and framework to communicate this new form of instructional leadership.

References

- Argyris, C. (1990). *Overcoming organizational defenses: Facilitating organizational learning*. Boston: Allyn and Bacon.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74.
- Burch, P. E., & Spillane, J. P. (2003). Elementary school leadership strategies and subject matter: Reforming mathematics and literacy instruction. *The Elementary School Journal*, 103(5), 519–535.
- Clay, M. M. (1985). *The early detection of reading difficulties: A diagnostic survey with recovery procedures* (3rd ed.). Auckland: Heinemann.
- Clay, M. M. (2000). *Running records for classroom teachers*. Auckland: Heinemann.
- Coburn, C. E. (2005). Shaping teacher sensemaking: School leaders and the enactment of reading policy. *Educational Policy*, 19(3), 476–509.
- Cunningham, P. M., & Allington, R. L. (1999). *Classrooms that work: They all can read and write* (2nd ed.). New York: Longman.
- Elmore, R. F. (2000). *Building a new structure for school leadership*. Washington DC: Albert Shanker Institute.
- Elmore, R., Peterson, P. L., & McCarthey, S. J. (1996). *Restructuring in the classroom: Teaching, learning, and school organization*. San Francisco: Jossey-Bass.
- Fountas, J., & Pinnell, G. (1996). *Guided reading: Good first teaching for all children*. Portsmouth, NH: Heinemann.
- Gamoran, A., Anderson, G. W., Quiroz, P. A., Secada, W. G., Williams, T. A., & Ashmann, S. (2003). *Transforming teaching in math and science: How school districts can support change*. New York: Teachers College Press.
- Greve, H. R. (2003) *Organizational learning from performance feedback: A behavioral perspective on innovation and change*. Cambridge, UK: Cambridge University Press.
- Hallinger, P., Bickman, L., & Davis, K. (1996). School context, principal leadership, and student reading achievement. *The Elementary School Journal*, 96, 527–549.
- Halverson, R. (2003, October 10). Systems of practice: How leaders use artifacts to create professional community in schools. *Education Policy Analysis Archives*, 11(37). Retrieved September 21, 2005, from <http://epaa.asu.edu/epaa/v11n37/>
- Halverson, R. (2004). Accessing, documenting and communicating the phronesis of school leadership practice. *American Journal of Education*, 111(1), 90–122.

Formative Feedback Systems and the New Instructional Leadership

- Halverson, R., Grigg, J., Prichett, R., & Thomas, C. (in press). The new instructional leadership: Creating data-driven instructional systems in schools. *Journal of School Leadership*.
- Halverson, R., & Thomas, C. (2007). *The roles and practices of student services staff as data-driven instructional leaders* (WCER Working Paper No. 2007-1). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved May 15, 2007, from http://www.wcer.wisc.edu/Publications/workingPapers/Working_Paper_No_2007_01.pdf
- Light, D., Honey, M., Heinze, C., Brunner, C., Wexler, D., Mandinach, E., et al. (2005). *Linking data and learning: The Grow Network study*. New York: EDC/Center for Children and Technology.
- Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, (64)4, 349–379.
- Ilgen, D. R., & Davis, C. A. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology: An International Review*, 49(3) 550–565.
- Kluger, A. N., & DiNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119, 254–284.
- Kreitner, R., & Kinicki, A. (2001). *Organizational behavior* (5th ed.). Chicago: Irwin McGraw Hill.
- Kruse, S., & Louis, K. S.. (2006, April). *Developing collective understanding over time: Building professional learning community top-down*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Lachat, M., & Smith, S. (2005). Practices that support data use in urban high schools. *Journal of Education for Students Placed at Risk* 10(3). 333-349.
- Louis, K.S. & Marks, H. (1998). Does professional community affect the classroom? Teachers' work and student experiences in restructuring schools. *American Journal of Education*, 106(4), 532–575.
- Mandinach, E. B., Honey, M., & Light, D. (2005, October). *A conceptual framework for data-driven decision making*. Paper prepared for the Wingspread Conference on Linking Data and Learning, Racine, WI.
- Murnane, R. J., Sharkey, N. S., & Boudett, K. P. (2005). Using student-assessment results to improve instruction: Lessons from a workshop. *Journal of Education for Students Placed At Risk*, 10(3), 269–280.
- Nadler, D. A. (1979). The effects of feedback on task group behavior: A review of the experimental research. *Organizational Behavior and Human Performance*, 23, 309–338.

Formative Feedback Systems and the New Instructional Leadership

- Peterson, P., McCarthy, S. J., & Elmore, R. F. (1995). Learning from school restructuring. *American Educational Research Journal*, 33(1), 119–154.
- Porter, A. (1995). *Critical issue: Integrating assessment and instruction in ways that support learning*. Retrieved May 15, 2007, from <http://www.ncrel.org/sdrs/areas/issues/methods/assment/as500.htm>
- Richardson, G. (1991). *Feedback thought in social science and systems theory*. Philadelphia: University of Pennsylvania Press.
- Senge, P. (1994). *The fifth discipline: The art and practice of the learning organization*. New York: Currency.
- Sharkey, N. S., & Murnane, R. J. (2006). Tough choices in designing a formative feedback system. *American Journal of Education*, 112, 572–581.
- Smith, M. S., & O'Day, J. (1991). Systemic school reform. In S. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp. 233–267). Philadelphia: Falmer Press.
- Smylie, M. (1994). Redesigning teachers' work: Connections to the classroom. *Review of Research in Education*, 20, 129–177.
- Spillane, J. P. (2000). Cognition and policy implementation: District policy-makers and the reform of mathematics education. *Cognition and Instruction*, 18(2), 141–179.
- Spillane, J. P. (2006). *Distributed leadership*. San Francisco: Jossey-Bass.
- Spillane, J. P., Halverson, R., & Diamond, J. B. (2004). Towards a theory of leadership practice: A distributed perspective. *Journal of Curriculum Studies*, 36(1), 3–34.
- Spillane, J., Reiser, B. J., & Reimer, T. (2002). Policy implementation and cognition: Reframing and refocusing implementation research. *Review of Educational Research*, 72(3), 387–431.
- Thorn, C. A. (2001, November 19). Knowledge management for educational information systems: What is the state of the field? *Education Policy Analysis Archives*, 9(47). Retrieved May 15, 2007, from <http://epaa.asu.edu/epaa/v9n47/>
- Von Bertalanffy, L. (1969). *General system theory: Foundations, development, applications*. New York: George Braziller.
- Wallace, M., & Pocklington, K. (2002). *Managing complex educational change: Large-scale reorganization of schools*. London: Routledge Falmer.
- Watson, J. G., & Mason, S. (2003, April). *Understanding schools' capacity to use data*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Formative Feedback Systems and the New Instructional Leadership

Wayman, J. C. (2005). Involving teachers in data-driven decision making: Using computer data systems to support teacher inquiry and reflection. *Journal of Education for Students Placed at Risk*, 10(3), 295–308.

Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.

Yeh, S. S. (2005, October 28). Limiting the unintended consequences of high-stakes testing. *Education Policy Analysis Archives*, 13(43). Retrieved May 15, 2007, from <http://epaa.asu.edu/epaa/v13n43/>