Routledge
Taylor & Francis Group

SECTION 1: MAPPING THE TERRAIN OF INTERIM ASSESSMENTS

School Formative Feedback Systems

Richard Halverson
University of Wisconsin-Madison

Data-driven instructional improvement relies on developing coherent systems that allow school staff to generate, interpret, and act upon quality formative information on students and school programs. This article offers a formative feedback system model that captures how school leaders and teachers structure artifacts and practices to create formative information flows across interventions, assessments, and actuation spaces. A formative feedback system model describes the organizational capacity upon which innovations such as comprehensive school reforms, benchmark assessment systems, and student behavior management systems draw to improve teaching and learning in schools.

This article presents the concept of a formative feedback system to identify the capacity that many schools are developing in the quest to meet the demands of high-stakes accountability policies. The press for raising student test scores has led many schools to reframe school instruction and information systems (Honey et al. 2005; Sharkey & Murnane, 2006; Wayman, 2005; Yeh, 2005). However, as Richard Elmore (2000) predicted, accountability-based reforms have called on schools to move beyond information system design to develop new kinds of capacity for instructional improvement. Because school staff cannot rely on standardized test results to directly inform changes in their classroom-level practices, schools must also engage in instructional system redesign—first to link everyday classroom practices with schoolwide outcomes, and second to develop data-driven practices that give teachers local, ongoing information to benchmark student learning progress.

Correspondence should be sent to Richard Halverson, University of Wisconsin-Madison, 1025 West Johnson 1161G, Madison, WI 53706. E-mail: halverson@education.wisc.edu

Our research on school capacity to use data for improving student learning has shown us how local actors develop data-driven instructional systems to improve classroom practice (Halverson, Grigg, Prichett, & Thomas, 2007; Halverson & Thomas, 2007; Prichett, 2007; Thomas, 2007). One key feature of this research has been to identify the operation of *formative feedback systems*. We proposed the phrase formative feedback systems to describe the networks of structures, people, and practices that help teachers and administrators translate testing data into practical information for everyday use (Halverson, Prichett, & Watson, 2007). Without accurate and timely formative feedback on the results of intended interventions, school leaders and teacher fly blind in their efforts to link what they expect to what actually happens in classrooms. Building these functions into the everyday school instructional program has proven to be a daunting, resource intensive challenge for school leaders. However, once assembled and implemented, formative feedback systems provide accurate, incremental, and actionable measures of student learning and behavior directly linked to the units of practice most meaningful to classroom teaching and learning.

This article presents an argument that formative feedback systems constitute sociotechnical processes that leaders and teachers coordinate to develop the capacity to make sense of assessment data. The article describes the structures and capacities that a selection of local school leaders developed to meet the information demands required by high-stakes accountability reforms. The phrase *formative feedback system* is used to name the local design work required to create the school capacity for meeting the demands of accountability policies. Our discussion uses some of the core ideas from the organizational change and information feedback literature to describe how these formative feedback practices operate and have emerged in the schools we studied.

The article focuses on data collected from four schools in order to illustrate a range of formative feedback systems in action. The case study analyses show how interventions can range from comprehensive school reforms to reading initiatives and student behavior recording systems, how assessments range from commercial products to locally developed benchmark assessment tools, and how teachers make sense of and act upon information in the context of daily teaching and learning. After identifying the key formative feedback system functions, the article presents an analysis of the intersection between the functions in order to explain where similar kinds of educational innovations fit into the family of formative feedback systems and to show how the characteristics of formative feedback systems can describe conditions for the successful implementation of benchmark assessments.

WHAT IS A FORMATIVE FEEDBACK SYSTEM?

Formative feedback systems draw on a rich body of research on learning and organizational theory. From a systems perspective, a school is a complex, messy information system that issues many conflicting performance signals (Wallace & Pocklington, 2002). Systems theory language helps to describe the organizational capacity to generate, interpret, and use feedback. Originally developed in cybernetics and engineering research (von Bertalanffy, 1969), feedback became a popular term in systems theory research (Greve, 2003; Richardson, 1991; Senge, 1994). Feedback is system-generated information that 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 (Richardson, 1991; von Bertalanffy, 1969). 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.

Formative feedback research in education has traditionally focused on the classroom. Black and Wiliam's (1998) work summarizes research on the critical role that formative feedback plays at the three levels of teacher, student, and teacher–student interaction. Formative feedback is information that can be used to guide both the teaching and learning process. At the teacher level, teachers need accurate information about the specific processes and outcomes of student learning to effectively shape 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. The policy dimension of formative feedback is designed to ensure that schools are able to tap into rich, generative bodies of formative information on the learning process and to make sure that the information is actually used by teachers and students to guide learning.

A school-level formative feedback system extends the insights from the classroom to the school as a learning organization. A formative feedback system model that would provide useful information about teaching and learning in schools would (a) generate information signals that measure how students performed in terms of an *intervention*, (b) develop sensor and processor functions to *assess* information signals, and (c) identify controllers that could *actuate* this new knowledge in order to adjust the instructional process. The three functions of *intervention*, *assessment*, and *actuation* compose the core components of a formative feedback system model (see Figure 1).

Interventions

Interventions describe the programs and policies that leaders and teachers use to guide student learning. At the group level, intervention artifacts can take the form of curricular materials, such as textbooks, experiments, worksheets, computer programs, and so on, that teachers use to structure classroom learning. At the individual level, a program such as the special education individualized education program (IEP) is an example of an intervention that customizes instructional and behavioral resources to meet the perceived needs of certain students in and out of the classroom. Taken together, these interventions compose the instructional program that the school staff uses to influence student learning. The learning that results from an intervention is analogous to the *signal* in classic information processing systems theory.

Assessments

Assessments pick up the information signal generated by the interventions. Assessments play the *sensor* role in the formative feedback system. Assessments provide the information to help teachers determine the degree to which signal received (estimates of student learning) correspond
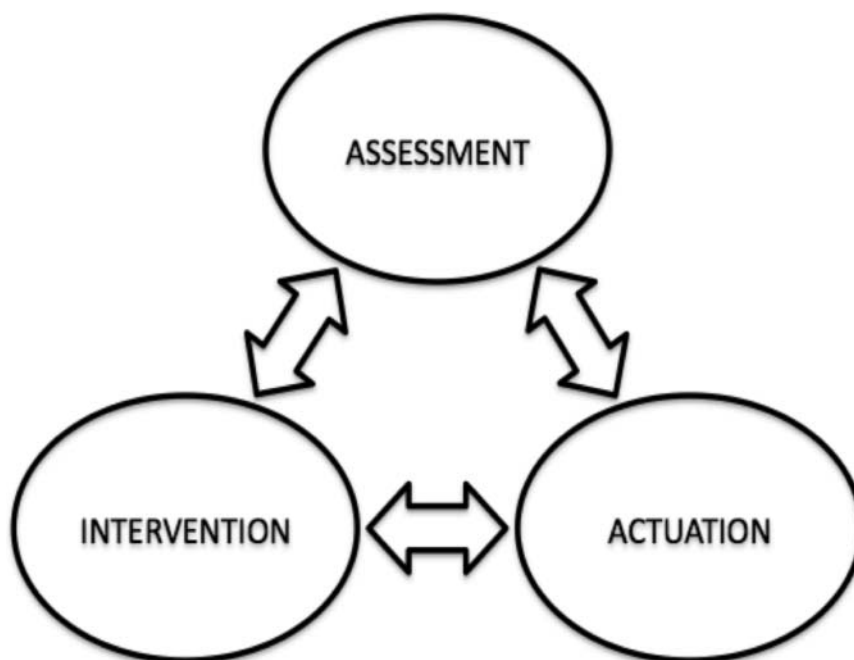
FIGURE 1   Formative feedback system model.

with the learning goals built into the interventions. Assessments provide specific information about the degree to which aspects of the intervention succeeded or failed to result in student learning. The match between the assessment and the intervention is critical—if the information generated by the intervention differs from the signal detected by the assessment, then the system will need a "translation" capacity to make sense of what the intervention results mean.

Actuation

Assessments of interventions, however well designed, merely provide information. Schools need structured occasions to turn assessment information into actionable knowledge. Actuation refers to 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 on the data and to make decisions about how to alter program delivery for students, or, in cases of significant problems revealed by the assessment, how to alter the interventions or the assessments themselves. Actuation spaces reflect the *signal processing* and the *controller aspects* of the classical feedback systems model. Actuation spaces provide legitimate occasions for practitioners to make sense of assessment information and, with adequate organizational support, provide the opportunity for staff to make appropriate adjustments to the intervention.

RESEARCH METHODS

Our study of *data-driven instructional systems* (DDIS; Halverson, Grigg, et al., 2007) investigated how school leaders and teachers are engaged in using data to redesign their local instructional and assessment practices. This article represents data collected during 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 instruction. In several of our schools, we found closed information subsystems that both generated information on student learning and program performance and provided legitimate contexts for faculty and staff to make sense of and act upon information. From this larger sample, we identified four schools that illustrate the design and operation of formative feedback systems (Table 1). This section reviews the methods we developed to conduct the DDIS research in general, and to describe the sites and artifacts we highlight in the findings and analysis sections that follow.

Data Collection

The study design documented data-driven leadership and instructional practices to describe the similarities and differences among schools' instantiations of the DDIS. To identify schools successful in using data to improve learning, 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. To document and describe the school-level DDIS we collected a variety of data, including 107 structured interviews with school teachers and leaders; 135 one-hr to three-hr observations of classroom teaching sessions, faculty meetings, professional development sessions, data retreats, and other important events as identified by the staff; and the collection of any documents that appeared relevant to data-driven practices, such as school improvement plans, staffing charts, budgetary information, and parent/community handouts.

TABLE 1
Formative Feedback System Data-Driven Instructional Systems Schools

School	Grades	Location	Size (Students)	Free/Reduced Lunch	Formative Feedback System	Principal Tenure
Pearson	K-6	Small town	300	60%	Balanced literacy	8 years
Walker	3–5	Rural	400	5%	Measures of Academic Progress	9 years
Malcolm	K-6	Urban	220	68%	Student behavior	6 years
Harrison	K-8	Urban	800	70%	Direct instruction	4 years

Data Analysis

The study approach to data analysis was a two-step, iterative process: (a) intrinsic case development (Stake, 2000) to construct 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) instrumental case development (Stake, 2000) to understand how this generic model works in individual schools. 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 six central functions for how successful schools use data to inform instruction (for a more detailed elaboration of the DDIS framework, see Halverson, Grigg, et al., 2007). These functions helped to 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.

  After constructing this initial model from across all of our case studies, we focused on the evidence of systemic formative feedback use to guide the coding of data within each individual school. We developed an iterative approach to a coding process (Strauss & Corbin, 1990) to sort our data into the DDIS categories and identified the key artifacts leaders and teachers used to address DDIS functions. We then reanalyzed the field notes, interviews, and documents to give us deeper insight into the individual schools' approaches to the DDIS. This article focuses on the data that emerged as relevant to understanding the schools' formative feedback systems.

FORMATIVE FEEDBACK SYSTEMS IN ACTION

DDIS researchers observed a variety of formative feedback practices throughout the classroom and professional interactions in our schools. Much of the feedback involved comments targeted toward individual students and concerned the direction or correction of student classroom work. Student-targeted feedback was communicated directly toward students; other feedback was shared among staff with the purpose of eventually influencing student learning. Feedback information seemed to occur as a natural outgrowth of teaching and learning, but a formative feedback system was something designed and maintained by school leaders. Although teachers generated and shared much of the information used as formative feedback, leaders took responsibility for establishing legitimate occasions (e.g., faculty meetings, role redefinitions, classroom assignments) that coordinated the flow of formative information. School leaders structured formative feedback systems as intentional efforts to coordinate information flow about performance quality across the instructional, assessment, and professional interactions spaces. Thus, although we observed a variety of formative feedback interaction in the schools we studied, we found relatively few formative feedback systems intentionally designed to elicit and use information to improve the instructional program as a whole.

  Picking out the specific functions from the tangled network of instructional and assessment practices in each school proved to be an ambiguous and messy task. Most school instructional practices evolve to satisfy multiple instructional, social, and personal functions. Simple interventions, such as scheduling changes or after-school programs, over time acquire new uses or become obstacles for new changes as they integrate with the organic development of a school culture. Artificially isolating information functions from the rich network of school practices runs the

risk of misrepresenting the degree to which the identified practices were intentionally designed to serve highlighted functions. However, the information functions identified as critical aspects of formative feedback systems did not appear to spring to life spontaneously from the vibrant stew of school culture. We found evidence of the intentional design of each of these information functions, and evidence that leaders and teachers intentionally coordinated these functions to provide formative feedback on central aspects of the school instructional program. Thus the findings presented here should not be taken as clean, abstract models to guide replication as best practices but rather as examples of how leaders and teachers created new forms of knowledge exchange in the midst of the exigencies of real school cultures.

The next section presents an analysis of formative feedback system functions that we identified across four of the DDIS schools. Each aspect of a school's formative feedback system appeared designed to elicit information about a specific aspect of the instructional program, and each involved the dedication of significant human and material resources to maintain a formative information flow. Although each of the schools in the study addressed aspects of all three components of the formative feedback model, two of the schools (Harrison and Pearson) were chosen to highlight the role of the intervention, one (Walker) to highlight assessment, and the Malcolm school case to highlight the role of the actuation space. The analysis illustrates the functions of formative feedback system components and how leaders and teachers designed features and linkages to ensure formative information flow.

INTERVENTIONS

The term *intervention* names the program or activities schools use to organize the school instructional program. Schools use a great variety of interventions to guide student learning, including curricula, student behavior programs, special education and guidance activities, and extracurricular activities. Each of these interventions provides structured or sequenced activities designed to influence student learning in some intended way. Halverson (2007) categorized interventions in terms of the *artifacts* that school staff use to influence student learning. Most of the interventions used for instruction in schools are received artifacts, that is, they originate from outside the school community. These artifacts included textbook and curriculum packages or computer systems. Districts often play a key role in selecting and distributing received intervention artifacts to schools. Locally designed artifacts include interventions such as teacher-assembled lessons, individualized education programs, and many after-school activities. Taken together, the aggregation of received and locally designed artifacts composes the school instructional program, a network of programs and policies that evolve over time as teachers and leaders add lessons, texts, and activities to their classroom instructional practices (Halverson, 2003). The following sections highlight two cases of instructional interventions—the Harrison comprehensive school reform and the locally designed Pearson literacy program—to illustrate how schools assemble programs and policies into instructional interventions that generate an actionable feedback signal.

The Harrison School Direct Instruction Program

A formative feedback system depends on the effort made by local school leaders and teachers to build an intervention that generates a coherent *signal* about the degree to which the

instructional program improves student learning. The comprehensive school reform program Direct Instruction (DI) played a central role as the intervention in Harrison School's formative feedback system. DI is a comprehensive school reform model that focuses on a tightly structured curriculum and assessment sequence. DI consists of a series of scripted lessons and orchestrated classroom interactions that provide both teachers and students with clear, prompt feedback on student learning. DI is one of a small number of comprehensive school reform programs found to have significant positive effects on student learning across implementations (Borman, Hewes, Overman, & Brown, 2003). Despite its predictable effects on student learning, DI has proven to be a controversial curriculum that has sparked critical debate about reducing teacher autonomy and decontextualizing student literacy skills from actual literacy practice (see, e.g., Commeyras, Shockley, Bisplinghoff, & Olson, 2003; Kuhn, 2007). Our interest here is not to engage in the debate over the pedagogical merits of DI but to note how the leaders and teachers at Harrison used DI to generate a reliable signal within a formative feedback system.

Harrison is an urban, culturally diverse K-8 school serving more than 500 students in a large Midwest city. Once identified as a "school in need of improvement" under the No Child Left Behind criteria, the Harrison staff applied for and received a Comprehensive School Reform grant to reorganize the school around the DI curriculum. DDIS researcher Chris Thomas's (2007) dissertation chronicled the story of Harrison's implementation of the Direct Instruction. The Harrison staff initially chose DI in the 1999 school year because the existing instructional program seemed to be a program in name only. To the incoming Harrison principal, it seemed as though "everybody was doing their own thing ... the former principal was not aware of a specific reading program ... and our reading resource teacher wasn't aware of a specific program either." The district played an interesting role in Harrison's selection of DI. In the early 2000s, the district approach to the school instructional program was to recommend the acceptance of common, districtwide learning outcomes and leave the selection of the means toward those outcomes to the local schools. The DI approach, in contrast, was to tightly link instructional means and outcomes and to remove local discretion from the instructional process. Further, there were significant differences between the district learning goals and the DI learning outcomes. Still, there were enough schools in Harrison's large urban district using DI that the district grudgingly recognized the DI schools and provided limited funding for district-level DI support.

Thomas (2007) related how the primary focus of the Harrison implementation of DI was for literacy in the early grades. The structured DI instructional program generated information about student learning in terms of a common curriculum and common learning standards. The signal produced by the DI instructional initiative comprised the performances of student learning that could be captured by the DI assessment system. In DI, the signal was coordinated due to the planned nature of the tasks engaged in by teachers throughout the school and across grade levels. These coordinated tasks allowed for the DI assessment to capture a coherent signal regarding student learning (in terms of the goals of the DI curriculum). The common approach to instruction allowed Harrison teachers to develop a shared vocabulary about instruction, and to discuss student learning in terms of a common curricular reference point. The shared DI curriculum enabled instructional staff to collaborate with teachers in analyzing the considerable amount of data produced in a typical DI classroom. Comparing the signal generated by the DI intervention enabled staff to identify which topics students were learning. Measuring the DI signal

with other kinds of assessments, such as DiBELS[1] and Six Traits Writing,[2] enabled the Harrison staff to determine the degree to which the DI instructional program needed to be supplemented with other curricula. DI thus provided the foundation for a common approach to instruction, and the Harrison staff were able to use the narrow range of the DI curriculum signal as an opportunity for the staff to determine which programs were needed to supplement the DI program.

The Pearson School Balanced Literacy Program

The Pearson School Balanced Literacy program provided a different approach to intervention assembly. The Pearson Elementary School is a small-town K-6 school of 300 students. The leaders and teachers at Pearson Elementary School organized a formative feedback system around early literacy instruction. Over the past 8 years, the Pearson principal and her teachers built structures that allow teachers to focus on particular learning problems, such as the development of early literacy skills, that limited student learning gains across the curriculum. In terms of our formative feedback system model, the reading curriculum was the *intervention*, a battery of commercial exams used by Pearson teachers was the *assessment*, and the regular grade level meetings for teacher reflection and action was the *actuation space*.

The Pearson Balanced Literacy program is an example of a locally designed instructional intervention. This package of curricula was stitched together under the guidance of the Pearson principal and Title I teacher. The impetus for the Pearson literacy program was a district (and state) press to improve the quality of K-2 reading teaching for all students. The building blocks for the Pearson Balanced Literacy program were Guided Reading, Reading Recovery, and Orton-Gillingham phonics. The Literacy Coordinator described that the Guided Reading program as "our Bible." Guided Reading leads students through a series of texts organized according to demonstrated reading levels. Trained Guided Reading teachers helped small groups of students use the contextual and visual cues of texts to understand the meaning of stories. Student grouping is determined by a series of quick assessments, or running records, that each teacher is expected to conduct to track student progress.

The Pearson Literacy Coordinator was trained as a Reading Recovery (RR) teacher, but the school found RR too expensive and intensive to serve the needs of all students at Pearson. RR was also limited by the recommended practice of pulling students out of the classrooms to receive services, which meant, according to the Literacy Coordinator, that "a child could have a different classroom teacher, reading teacher, and Reading Recovery teacher." She found RR an "absolutely excellent" complement to Guided Reading, and she worked with the K-2 teachers adapt some of the RR writing materials into the literacy program. The Pearson staff found that even with the RR supports, the Guided Reading program proved difficult to implement with new readers who struggled with simple phonics skills. Pearson reading teachers began to draw on Orton-Gillingham

---

[1]DiBELS, or Dynamic Indicators of Basic Early Literacy Skills, are a set of standardized, individually administered measures of early literacy development. They are designed to be short (1 min) fluency measures used to regularly monitor the development of prereading and early reading skills (http://dibels.uoregon.edu/).

[2]Six Traits of Writing is a comprehensive intervention developed by the Northwest Regional Education Laboratory that organizes the writing process in terms of ideas, organization, voice, sentence fluency, and word choice (http://www.nwrel.org/assessment/department.php?d = 1).

techniques to enhance the phonics aspects of the balanced literacy program. Orton-Gillingham involved daily practice in sounds and word decoding skills to prepare students for book reading. The staff began to use Guided Reading sessions to as diagnostic sessions to identify the kinds of phonics support students needed. The common staff training and commitment to Guided Reading, Reading Recovery, and Orton-Gillingham helped provide a shared vocabulary for instruction so that "now the language is common, and even our struggling readers understand what we are talking about." Teachers at Pearson credited the effective use of Balanced Literacy generated 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."

The role of an intervention in a formative feedback system is to transmit a clear signal about student learning that can be measured through assessments and acted upon in actuation. In practice, the focus on producing an actionable signal to indicate student learning meant that the Pearson staff needed to standardize and streamline their locally assembled curricular intervention to produce a clear signal. The need to coordinate interventions to produce a clear signal had important implications for school program design. Leaders and teachers at Harrison and Pearson had to collaboratively agree to standardize teaching across the grade level by coordinating the content and pacing of their instruction. If teachers would select their own curricular activities, or if they would decide on their own pacing, then it would be difficult to ascertain a clear signal that could be readily compared across classrooms or grades to report student learning progress. The coordination of the instructional program across classrooms serves as a boundary definition for whether the school had developed a formative feedback system signal in a given instructional area.

The difference between adopting received or locally designed intervention artifacts led to distinctive, but ultimately convergent, formative feedback system challenges. At Harrison, the decision to adopt Direct Instruction required teachers to commit to common instructional program selection and a shared pacing schedule. As a result, the literacy intervention produced a clear signal about student learning. However, the challenge with DI at Harrison was to discern the *adequacy* of the signal as an indicator of student learning. The state test score results showed that the Harrison DI curriculum led to student learning gains in certain domains, but it left gaps in student performance in other areas (e.g., reading comprehension in the upper grades, and writing). Comparing the Harrison signal with the state test score results called into question the adequacy of the DI intervention signal to indicate the scope of intended student learning outcomes. The Harrison formative feedback system needed to investigate the how to supplement the instructional intervention in order to address student needs across learning domains. The locally designed intervention at Pearson faced a different challenge—signal *coherence*. Each component of the Balanced Literacy program (Guided Reading, Reading Recovery, Orton-Gillingham phonics) produced a different kind of signal. Pearson staff had to collectively interpret what each signal meant in terms of student learning and to decide on practices that would coordinate information to elicit a clear signal for each student. Ultimately, the relation between signal coherence and adequacy ought to spark an iterative design process in a healthy formative feedback system. At Harrison, the addition of supplemental programs to address signal adequacy issues raised the problem of how the programs fit together as a signal coherence issue, and after building the Balanced Literacy program, Pearson staff had to compare the resulting signal with the independent state test results to again gauge signal adequacy. Thus the establishment of a clear intervention signal can be seen as both a condition for and an ongoing process within a vibrant formative feedback system.

ASSESSMENTS

The role of an assessment in a formative feedback system is to sense the signal sent by the intervention in a way that facilitates corrective action by school staff. A formative feedback system assessment translates the intervention signal into information that helps to create a shared understanding of how staff should act to improve student learning. Schools are typically awash in many different types of assessment data, and it is a significant design challenge to constrain the various assessments to produce a reliable and shared measure of student learning (Hamilton et al., 2007). School learning assessments exist at many different levels and serve different purposes. Schools and districts are held accountable by government for documenting student learning in terms of summative standardized tests. Local schools and classrooms receive and design a wide variety of formative assessments, ranging from benchmark assessment systems, to teacher developed quizzes and homework checks, to monitor the learning process. A challenge of formative feedback system design is to establish a direct link between interventions and assessments to create actionable information for faculty and staff.

Benchmark assessment systems have recently emerged as tools for schools to coordinate assessment information. Benchmark assessments aim to provide timely and appropriate data to guide schools in making effective decisions about teaching and learning (Burch, 2010/this issue; Li, Marion, Perie, & Gong, 2010/this issue). The systems typically involve output processes to deliver the assessment information in student-level or learning-standard-level reports that make sense for guiding teaching and learning. Some benchmark systems are computer adaptive tests that narrow the range of appropriate items offered to individual test-takers based on responses to prior items (Cronin, 2004). Items are selected according to the match with state or local standards in math, literacy, and science. Benchmark assessment tools can be either locally designed or comprehensive, proprietary systems. Proprietary benchmark assessment systems, such as the Northwest Evaluation Association's Measures of Academic Progress (MAP) or McGraw-Hill's *Acuity*, are typically implemented as district-wide assessment products that aim to provide ongoing information on student progress toward learning goals. Computer-based benchmark systems also address the time lag involved in standardized testing—results are typically available to schools in less than 3 days. Finally, many benchmark assessment tools provide reporting tools to convey the results of testing in terms teachers can act upon, and curriculum tools to point teachers toward viable course of instructional action (Prichett, 2007).

The Walker school experience with implementing the MAP benchmark assessment system provides a good example of the design challenges involved in implementing a benchmark system tool to play an assessment role in a formative feedback system. Walker is a Midwest rural Grade 3 to 5 school with an enrollment of about 400 students. About 90% of the Walker students score proficient or advanced on the state test. Reid Pritchett's (2007) dissertation work related the story of how the Walker school leaders and teachers worked to create formative feedback assessment capacity by implementing MAP. In 2004, the Walker district decided to purchase MAP and worked with Walker teachers to integrate MAP into the instructional program. The Walker implementation of MAP was initiated by district leaders eager to acquire the capacity for systemwide, intermittent measures of student learning. The district curriculum director perceived that the state test did not generate sufficiently actionable information, and, by 2003, "(we) didn't have an assessment tool where we could get information quickly about reading, writing, and math. And we needed not only for something to be easily administered, but also get the information

quickly back." In 2004, the district developed a committee of leaders and teachers who decided to purchase and implement MAP in Grades 3 to 8. The committee felt confident that MAP would provide the kinds of information district leaders thought teachers would need to make adjustments in instructional practices. In 2004–2005, the district purchased MAP and helped district schools coordinate the professional development sessions necessary to help teachers integrate MAP data into their classes.

The Walker school principal clearly saw the possibilities of MAP for classroom formative feedback. She thought that MAP would help teachers to measure students in terms of local standards, to "give (teachers) a kind of a foundation . . . a consistent way of analyzing where a student might fall within expectations, either by the state or their national norms, and then how that aligns with what they're doing in the classroom." Further, MAP would provide a context for "a consistent way of assessing a child and then to reflect on how consistent that is with what they're seeing in their individual classrooms." The Walker teachers, however, were hesitant about the possibility of MAP being used as a teacher evaluation tool. Prior to purchasing MAP, Walker teachers had already developed a basic formative feedback system (similar to the Pearson school model) in literacy and language arts, and some teachers regarding MAP implementation as an opportunity to coordinate the math instructional program. Other teachers, however, saw different purposes in the district-initiated MAP implementation. One teacher commented that MAP data could add to whether "they decide they want to get rid of you, that's one thing that they could use against you." Over time, teacher concerns about teacher evaluation seemed to ease as the principal made a significant time commitment to help teachers make sense of the MAP data reports in terms of math instruction. The Walker principal used MAP data in faculty and staff meetings to create a common vocabulary for Walker teachers to discuss student learning. After the 2nd year of MAP implementation, one teacher noted that MAP provided an "extra set of eyes to confirm that you're doing something well or that you need to do something different."

The experience of implementing MAP at the Walker school reflected a tension between purchased benchmark assessment systems and a functional, schoolwide formative feedback system. Walker's initial implementation of MAP focused on how teachers could individually use MAP data to improve learning. After 3 years, 80% of teachers reported using MAP data to group students in their own classrooms, but only 30% reported using MAP to evaluate the effectiveness of the instructional program (Prichett, 2007). In a schoolwide formative feedback system, leaders and teachers use information to make adjustments not only for individual students but also to the interventions themselves. The principal began to use MAP assessment reports to justify conversations with each teacher about learning with each student in their classrooms. These conversations helped teachers to see how the MAP data could be used to guide instructional improvement. As a result of these conversations, after 3 years Walker teachers were beginning to create linkages between MAP data and day-to-day classroom lessons. In math, for example, a team of teachers from across grade levels painstakingly analyzed district geometry curricula to determine just where the MAP assessments of student learning linked with the district curriculum that guided daily classroom instruction.

The Walker school implementation of MAP illustrates some of the issues that arise in use of widely available assessment tools to create a functional formative feedback system. Formative feedback systems depend on the ability of the assessment sensor to interpret the intervention signal in a format that can directly support teacher action. At Walker, because the MAP assessment and the curriculum are both aligned with state standards, school leaders, and teachers, one might have

assumed that the assessment and the intervention would be aligned as well. The Walker teacher experience suggests that the mere alignment of signal and sensor will not produce actionable system-wide information. The key for understanding the difficulty of using MAP in a formative feedback system is that teachers teach lessons, not standards, and that assessment tools such as MAP are aligned to standards, not lessons. Porter (1995) argued that unless there is a tight match between what is assessed and what is taught, the assessment results can be 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. The key assessment design issue in a formative feedback system lies in creating an actionable fit between intervention signal and the assessment sensor. Teachers who put the time into understanding the relation between, for example, the MAP lexile scores and actual reading comprehension lessons, may be able to use MAP data to guide instruction. Absent efforts to collaboratively build connections between day-to-day interventions and assessments, benchmark assessment systems may simply reinforce the isolated expertise of teachers willing to create formative feedback from benchmark data.

ACTUATION

Creating opportunities for teachers and staff to consider and discuss assessment data 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). Each school provided examples such opportunities, or *actuation spaces*, such as grade-level and faculty meetings in which teachers reflected on student learning and behavior data. However, only a few of the actuation spaces we observed appeared to be directly connected to the interventions and assessments of a formative feedback system. Three features seemed to differentiate formative feedback actuation spaces from ordinary meetings. First, the agenda for the actuation spaces were dominated by ongoing conversations about assessment data. Second, staff included in the actuation meetings took persistent roles as both designers and implementers of instructional interventions and assessments. Team members felt authorized to use assessment data for taking action at both the student and the intervention level. Third, school leaders played a key role in scheduling adequate time and resources for actuation spaces, negotiating agendas and the range of responsibilities with team members, and distributing authority so that the actuation teams were empowered to take the action deemed necessary. The design and implementation of actuation spaces showed how the social distribution of leadership unfolded in practice (Spillane, 2006).

The Malcolm School Respect and Responsibility (R&R) program provides an example of how leaders structured actuation space within a formative feedback system. Malcolm is an urban Midwest K-5 school with a highly mobile population of 220 children. In the past several years, 70% of Malcolm's students have qualified for free or reduced lunch. The school's proportion of minority students is among the highest in the district. Malcolm is a schoolwide Title I school, and it is eligible for state class-size reduction funding. Malcolm had improved their student standardized test scores to the point that they rivaled other district schools. Leaders at Malcolm felt that the learning issues of the predominantly low-income student population could best be addressed with a good curriculum and a safe learning environment. The Malcolm district provided nationally renowned curricular initiatives in reading and math that included common lesson design across

classrooms, benchmark subject area assessments, and guidelines for faculty collaboration and action. Beginning in 2002, Malcolm school leaders decided to focus their efforts to locally design a program to monitor and manage student behavior to create a safe learning environment.

Malcolm R&R program was a data-driven support system for student behavior that addressed both program-level and student-level concerns at the school. The student service staff, including school psychologists, Title I teacher, special educators, and social workers, facilitated these supports as well as played an integral role in making sense of data taken on program and student level interventions. For Malcolm school, a variety of interventions constituted the intervention; the student behavior information system was the Assessment; and the R&R, grade-level, and faculty committees provided the actuation spaces for the formative feedback system. The school developed R&R from a number of traditional policies and programs intended to help students develop conflict management strategies and to guide teachers in addressing behavioral issues in the classroom. R&R data were gathered through multiple reporting tools and compiled in a networked spreadsheet that records incidents reported on a standard form including the nature of the incident, the time and location, and the prior interventions attempted.

R&R served as a two-stage actuation space that brought faculty together for regular discussions about the behavior and learning patterns for individual students. The first state was composed of the R&R team. Student behavioral and learning data were reviewed weekly by the R&R committee, a group of nonteaching staff who regularly met to monitor and reflect upon the student data in the hope of identifying emergent systemic issues within the school. Some issues seemed to be behavioral, such as the high frequency of referrals as students return from the playground, whereas others concerned instruction, such as the frequency of special education and learning problem referrals. The R&R committee compiled a monthly report that was shared with staff at grade-level meetings. The grade-level meetings constituted a second stage of the actuation space in which all faculty members would address questions such as the persistence of data across years (for groups or for individual students), the distribution of behavioral and academic referrals by grade, and referral trends for individual students. One member of the R&R team offered an example of how their data review process led to an intervention on behalf of a teacher: "We helped the teacher to change the layout—the physical plans, where the desks were, where the work space was and got rid of a lot of clutter. Then the referrals (for her classroom) went down." The principal appeared at each grade-level meeting to share the data with them and to identify any patterns—positive or negative—that may arise. As the principal noted, the more frequent review of the data has allowed them to be "much more aggressive with interventions."

Malcolm leaders provided another level of actuation space to address student issues that went beyond the scope of R&R—the Building Consultation Team (BCT). A BCT was convened to address acute issues pertaining to a particular student. Similar to the special education IEP process, a BCT committee was convened to address behavioral issues for individual students (see Halverson & Thomas, 2007, and Thomas, 2007, for further discussion). The BCT includes staff members who interact with students in a number of different contexts, including the classroom teacher, support staff, and school administrators; parents are always invited to the meetings and frequently attend them. The BCT was seen as a preliminary step before the school would develop an IEP. Some BCT students ultimately ended up being referred for an IEP, but many did not. Sharing and responding to the data provides a purpose for its collection and, in the case of this school, leads to increased demand for it. As a district administrator commented, "You have got to use data. I mean, you're not just creating data to create data, you have to create it for a

purpose," and the school's principal told her staff, "The beauty of data is that we can have these conversations."

The R&R actuation spaces illustrate several key features of formative feedback systems. First, the actuation spaces legitimate a schoolwide *sense-making* process. Sense-making reflects individual and group efforts to notice or interpret new stimuli in light of prior knowledge, belief, and experience (Spillane, Reiser, & Reimer, 2002). Actuation spaces create contexts in which staff can engage in collective sense-making to make decisions about teaching and learning. The school's principal once remarked that "there's not one secret with data in this building." The R&R actuation spaces scaffold the sense-making process for Malcolm staff. The R&R committee sifted through behavior reporting system data to discern teacher and student patterns. Rather than hold staff accountable for these findings, the second-stage actuation space—grade-level meetings—are structured to provide opportunities for teachers to reflect on what the behavioral data reports mean in terms of teacher practice. Transforming faculty meeting times into structured sense-making spaces legitimates the data collection process and creates an expectation that data review sessions are a necessary aspect of professional practice at Malcolm.

Second, the actuation spaces create the capacity for schools to use appropriately. generated information to alter interventions—either at for individual or for the organization as a whole. The link between actuation spaces and intervention represents 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. The BCT sessions often follow directly from the R&R meetings to build a learning or behavioral plan for an individual student. The grade-level meetings encourage teachers to discuss changes in classroom practices. The Malcolm actuation spaces also create the perceived need for alteration or design of schoolwide interventions. For example, in 2005–06, the erosion of student behavior led Malcolm staff to design a series of academic parent–community nights designed to improve family understanding of district curricular programs. The staff felt that if parents could work with students to improve academic performance, students would be less frustrated in school and more willing to engage in classroom lessons. This function of actuation spaces reinforces the legitimacy of the entire formative feedback system. If teachers and staff generate, record, and analyze data to develop or refine existing interventions, only to be reminded that they do not have the power to alter the instructional program, subsequent cycles of the feedback system are much less likely to be formative. The dense interaction around behavioral data at Malcolm (as with literacy data at Pearson and Harrison) appeared to contribute to a vibrant professional community in which teachers felt comfortable making their classroom practice public. The sustained success of a formative feedback system in a loosely coupled school system appears, at least in part, to rest on the belief that the actions taken as a result of the data will count in shaping future practice.

CONCLUSION

This article argued that the concept of formative feedback systems describes an important aspect of data-driven instructional capacity in schools. A systems theory approach enabled the analysis of school data practices around the language of signals, sensors, processors, and actuation. Formative feedback systems are composed of links between interventions, assessments, and actuation spaces

that enable school staff to act upon student learning information. A formative feedback system model can thus be seen as a kind of organizational capacity designed to develop and distribute information about the instructional program. Seen in this light, a formative feedback system provides a generic description of the capacity that tools such as comprehensive school reforms, benchmark assessment systems, or student behavior management programs seek to cultivate in schools. Such interventions rely on the development of interventions, assessments, and actuation spaces, and on the designed interaction among these components, to act as effective formative feedback systems.

The argument was designed to make a modest claim about how leaders and teachers build capacity for data-driven instructional practices in schools. Neither the concept nor the programs identified as components of formative feedback systems are new. Thousands of schools use phonics, Guided Reading, running records, faculty meetings, and behavior management programs to conduct their business. The cases presented from the DDIS study illustrate how school data practices might be analyzed collectively as the result of design in terms of formative feedback system functions. School leaders and teachers in some of the DDIS schools used humble components to construct powerful formative feedback systems that used sophisticated, carefully selected information to enable program customization on the fly. Describing a model that captures these designed feedback loops of instruction, assessment, and actuation may lead researchers to 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 organizations. 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

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, *5*(1), 7–74.

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research*, *73*, 125–230.

Burch, P. (2010/this issue). The bigger picture: Institutional perspectives on interim assessment technologies. *Peabody Journal of Education*, *85*(2), 147–162.

Coburn, C. E. (2005). Shaping teacher sensemaking: School leaders and the enactment of reading policy. *Educational Policy*, *19*, 476–509.

Commeyras, M., Shockley, B., Bisplinghoff, M., & Olson, J. (2003) *Teachers as readers: Perspectives on the importance of reading in teachers' classrooms and lives*. Newark, DE: International Reading Association.

Cronin, J. (2004) *Aligning the NWEA RIT scale with the Wisconsin Knowledge and Concepts Exam*. Lake Oswego, OR: Northwest Evaluation Association. Retrieved September 12, 2007, from http://nwea.org

Elmore, R. F. (2000). *Building a new structure for school leadership*. Washington, DC: Albert Shanker Institute.

Greve, H. R. (2003) *Organizational learning from performance feedback: A behavioral perspective on innovation and change*. Cambridge, UK: Cambridge University Press.

Halverson, R. (2003). Systems of practice: How leaders use artifacts to create professional community in schools. *Education Policy Analysis Archives*, *11*(37), 1–35.

Halverson, R. (2007). A distributed leadership perspective on how leaders use artifacts to create professional community in schools. In L. Stoll & K. S. Louis (Eds.), *Professional learning communities: Divergence, detail and difficulties* (pp. 93–105). Maidenhead, UK: Open University Press.

Halverson, R., Grigg, J., Prichett, R., & Thomas, C. (2007). The new instructional leadership: Creating data-driven instructional systems in schools. *Journal of School Leadership*, *17*, 159–193.

Halverson, R., Prichett, R. P., & Watson, J. G. (2007) *Formative feedback and the new instructional leadership* (Wisconsin Center for Educational Research Working Paper 2007–3). Madison: Wisconsin Center for Educational Research. Retrieved January 17, 2009, from http://www.wcer.wisc.edu/publications/workingPapers/Working_Paper_No_2007_03.pdf

Halverson, R., & Thomas, C. (2007). The roles and practices of student services staff as data-driven instructional leaders (Wisconsin Center for Educational Research Working Paper 2007–001). Madison: Wisconsin Center for Educational Research. Retrieved February 13, 2007, from http://www.wcer.wisc.edu/Publications/workingPapers/Working_Paper_No_2007_01.pdf

Hamilton, L. S., Stecher, B. M., Marsh, J. A., McCombs, J. S., Robyn, A., Russell, J. L., et al. (2007). *Standards-based accountability under No Child Left Behind experiences of teachers and administrators in three states*. Santa Monica, CA: RAND Corporation

Honey, M., Brunner, C., Light, D., Kim, C., McDermott, M., Heinze, C., et al. (2005). *Linking data and learning: The Grow Network study*. New York: EDC/Center for Children and Technology.

Kuhn, D. (2007). Is direct instruction an answer to the right question? *Educational Psychologist*, *42*(2), 1–5.

Li, Y., Marion, S., Perie, M., & Gong, B. (2010/this issue). An approach for evaluating the technical quality of interim assessments. *Peabody Journal of Education*, *85*(2), 163–185.

Porter, A. (1995). *Critical issue: Integrating assessment and instruction in ways that support learning*. Retrieved on October 30, 2006, from North Central Regional Educational Laboratory Web site: http://www.ncrel.org/sdrs/areas/issues/methods/assment/as500.htm

Prichett, R. (2007). *How school leaders make sense of and use formative feedback systems*. Unpublished doctoral dissertation, University of Wisconsin-Madison School of Education, Madison, Wisconsin.

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.

Spillane, J. (2006). *Distributed leadership*. San Francisco, CA: Jossey-Bass.

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.

Stake, R. E. (2000). Case studies. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 435–454). Thousand Oaks, CA: Sage.

Strauss, A., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory, procedures and techniques*. Newbury Park, CA: Sage.

Thomas, C. (2007). *Problem-solving teams and data-driven school leadership: A path toward the next generation of special education services*. Unpublished doctoral dissertation, University of Wisconsin-Madison School of Education. Madison, Wisconsin.

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.

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.

Yeh, S. S. (2005). Limiting the unintended consequences of high-stakes testing. *Education Policy Analysis Archives*, *13*(43). Retrieved February 12, 2007, from http://epaa.asu.edu/epaa/v13n43/