

Use-inspired Research and Development: Solving Middle School Challenges in the SERP-Boston Field Site

Abstract

In the field sites of the Strategic Education Research Partnership (SERP), district partners define the agenda for an interdisciplinary team of researchers, developers and practitioners. Researchers from Harvard University, Wheelock College, and ETS along with Boston Public Schools collaborators describe how multiple lines of interrelated work on instruction, assessment and school organization were conducted in schools to address complex middle school literacy challenges. Presenters focus a lens on how core principles of the partnership model shaped the work's evolution and stakeholders' ability to bridge boundaries and create synergy. As the partnership approaches its five year milestone, lessons from this productive collaboration can inform the work of others who share the mission of conducting "use-inspired" research and development in "Pasteur's Quadrant."

Session Summary

Suzanne Donovan
SERP Institute

Objectives: In 2003 a National Research Council report made the case for an education research and development enterprise conducted in "Pasteur's Quadrant," where research is done in the interest of solving problems of practice. The report called for development of school district "field sites," hypothesizing that locating the work in practice settings would change its direction, content, and use value. SERP has been engaged in such collaboration with the Boston Public Schools (BPS) for five years. This session will present the interdisciplinary program of work under way, and the profound nature of the partnership's influence on its evolution.

Overview of the presentations: At the District's request, the partnership first turned its attention to the urgent problem of students entering high school unable to comprehend their textbooks, and targeted middle school instruction for the intervention. Given the lack of systematic attention to the academic vocabulary students need to comprehend discipline-specific texts, Catherine Snow directed a SERP team that developed and piloted *Word Generation*. Content area teachers were integrated into the program to help ensure adequate exposure to target words in multiple contexts, while requiring teachers to cross subject area boundaries—a major challenge for schools with low internal coherence. A second line of work was formulated to address the issue of low internal coherence. Led by Richard Elmore, it had two purposes: first, to develop an instrument capable of capturing the coherence level in a school and second, to develop protocols for building coherence that could be used in conjunction with *Word Generation* or other school improvement efforts. The district's need for a fine grain diagnostic instrument resulted in a third line of work with the goal of matching struggling readers to appropriate programmatic options. A SERP team led by John Sabatini designed the *RISE* assessment

to generate student reading and comprehension profiles. Results from piloting the *RISE* indicate that a large segment of the student population still struggles with basic reading skills that are not being taught in middle school. In response to this latter finding, a fourth line of work was initiated and led by Lowry Hemphill who is developing and piloting the *Strategic Adolescent Reading Intervention* with components targeting teacher preparation as well as student instruction.

Significance: The dominant model for education research is one in which the motivating questions, theories, and approaches to reform are rooted in the university. In contrast, the SERP partnership model demonstrates that the context in which work transpires substantially influences the work itself.

Symposium structure: Following an overview describing the school district context, four lines of work will be presented in sequence. Inspired by the conference theme, the participants will emphasize a district's complex environment and how programs interact within it. They will illustrate the ways that the context of a partnership fundamentally changes the character and content of its work. Three expert discussants including two literacy leaders from BPS will then discuss the contributions and challenges of integrating research and practice from a district perspective.

Building Curriculum in Response to District Need

Catherine Snow

Harvard University

Objective. *Word Generation* is a middle-school vocabulary and academic language curriculum designed in response to the Boston Public Schools' (BPS) desire to improve students' literacy skills.

Perspectives. The program design has two major influences: a) research on vocabulary development and b) observations revealing domains where many BPS teachers were seeking help. These domains included teaching the vocabulary encountered in student texts, promoting deep comprehension of nonliterary texts, and guiding productive classroom discussions.

Methods. The program was organized around issues selected by a focus group of BPS practitioners as likely to be of interest to middle schoolers, reflecting a commitment to providing activities and topics that would engage young adolescents. Activities designed for the language arts, math, science, and social studies teachers to implement were vetted and modified by the teachers themselves, most intensively during the first year of the program when teacher-researcher collaboration was most intensive. Thus, the program has been carefully adapted to the needs of the teachers and students in classrooms where it has been used. The effectiveness of the program was monitored using student improvement from pre- to post-test scores, as well as evidence of student engagement and teacher satisfaction.

Evidence/Results. Test results after 12 weeks of *Word Generation* participation showed improvement among students on the curriculum-based test that was equivalent to approximately two years of incidental vocabulary growth (Snow, Lawrence & White, in press). At the same time, the effect size varied across schools from approximately .2 to as much as .6. Those school-level differences in effect were associated with variation in the quality and thoroughness of the implementation of *Word Generation* across the schools. The SERP context made it possible to collaborate with the Internal Coherence team to explore the degree to which the schools with smaller effect sizes were lacking internal accountability and strong instructional leadership. Precisely because this program relies on the collaboration of teachers across subject areas, the presence of structures (grade-level team meetings, shared professional development) focusing teachers on accountability for student learning may be crucial to its success. Schools with low internal coherence were unlikely to use the program as it was designed—as a whole-school, student-focused effort. As a result, *Word Generation* has been linked to efforts to promote internal coherence at the school level by other SERP team members.

Significance. The *Word Generation* program is a model with proven effectiveness in improving middle-school student outcomes; its success depends to a large extent on the contribution to its design of practitioners who were actually using the program, figuring out which topics and which activities engaged students, and how the materials could be designed for optimal access. At the same time, the quality of implementation depended, not just on the collaboratively produced usability of the program, but also on features of the schools in which it was being used—demonstrating the importance of thinking about curricular and school-level organizational issues together in order to improve student outcomes.

Building Knowledge and Supports for Improving School Coherence

Richard Elmore

Harvard University

Objective: *Internal Coherence* (IC) is a measure of a school's ability to function as a unified organization and to run effective organizational processes in pursuit of collectively determined goals. The SERP IC component was developed to act in concert with the implementation of *Word Generation*, the vocabulary and academic language curriculum targeting middle-school literacy, in recognition of the fact that even the most powerful intervention introduced into an atomized organization will not have the desired whole-school effect on student learning. The SERP IC initiative focuses on building schools' organizational capacity to improve instruction across classrooms, from one year to the next, by targeting classroom-, school- and district-leadership.

Perspectives: The IC project is grounded in research from several domains, including organizational learning, internal accountability, collective efficacy, and distributed leadership.

Method: For the past three years, IC researchers and Boston school district administrators have worked on a model for assessing the capacity of schools to engage in sustained, school wide, instructional improvement. IC researchers developed a diagnostic process for measuring the organizational capacity of schools and piloted the process in a sample 21 low-performing schools in Boston. The diagnostic consists of a teacher survey, which draws on existing research and measurement practices, as well as a series of corroborating interviews and classroom observations. In addition to the diagnostic, the IC project has also engaged in qualitative research in a sample of schools to document practitioners' initial responses to receiving their assessment data, and has engaged in work with system-level and school-level leaders on how to use schools' assessment data to contribute to their school improvement efforts.

Evidence/Results: We have done basic statistical analyses to examine the degree to which the items in the assessment survey measure discrete dimensions of organizational capacity, which confirmed their validity, and basic analyses of variations across the sample of 21 schools on the survey dimensions. We have also analyzed the interview and observation data for similarities and differences, by school, in the patterns that the survey presents. One preliminary finding was that there was much more variability among schools on organizational capacity measures than we expected, and this pattern of variation was confirmed by the qualitative interview data. This finding suggests that there are some low-performing schools that actually have the capacity to make significant improvements in instructional quality and performance, but are "stuck" at a particular performance level and need guidance on the next level of work. Other schools are low-performing and score relatively low on organizational capacity, suggesting that they need either considerable work or reconstitution.

Significance: The advent of performance-based accountability in education, beginning in the early 1980s and reaching its peak with the passage of No Child Left Behind (NCLB) in 2001, has put many schools and school systems under unprecedented pressure to make systematic, sustained gains in measured performance over time. Despite this fact, many schools remain unable to manage their internal organizations in ways that lead to coherent improvement in student performance over time.

Designing and Implementing School Level Assessments with District Input

John Sabatini

Educational Testing Service

Objective: The *RISE* was conceptualized as a consequence of design team meetings with district leadership. In those meetings, we learned that district leaders had made significant investments in reading intervention products (e.g., Read 180) and in teacher professional development to support literacy. However, they had no consistent reading assessment systems in schools, other than student state test scores, to provide guidance as to the nature and severity of student difficulties with reading. In some schools, traditional school level reading comprehension tests were given, but these largely duplicated the

score results of the state test, affirming that struggling readers were indeed struggling. Thus, their efficiency in aligning reading instruction to student needs was lacking.

Perspectives: *RISE* prototypes were designed to be aligned with current empirical research on the nature of struggling reader difficulties and effective instructional programs. Specifically, the *RISE* targets word identification, reading fluency, vocabulary, and basic comprehension processes. Prototypes have included both constructed response items such as oral reading tasks, and choice response formats.

Methods/Techniques: Working directly with districts and school personnel revealed other complexities and constraints that shaped iterative *RISE* prototypes. First, in order for the information to be actionable, school leaders needed results early in the school year, so rapid turnaround of scores was required. This led to designing a computerized battery with all choice responses, which could be scored automatically. Second, despite the lack of diagnostic information on students, the reality was that middle school students were over-tested across the year. There was a desire, therefore, to minimize the amount of additional testing. This led to the decision to encapsulate the six subtest battery into one 45-50 minute session. Third, there was a desire that the content of tests be relevant to the district curriculum. This led to the decision to model content on curriculum materials that middle grades students were likely to encounter in different subject areas, and to benchmark subtest scores to state test results.

An ongoing design issue concerns the heterogeneity of school technology capacity and security. A software website download to each school computer avoided problems of web delivery such as unreliable connectivity and bandwidth limitations. However, as we began implementation in schools, we learned that software in some schools automatically erased any data saved each day on the computers. This led to a host of adaptations tailored to circumvent the problem.

Evidence/Results: The *RISE* battery has been successfully implemented in progressive years in middle school and high school settings to over a thousand students. Reliability and other test properties have improved with each iterative design, while the entire battery continues to be administered within a 50 minute session. Schools have found results useful and continue to support us in enhancing the quality and features of the *RISE*. Data supporting these claims will be reported.

Significance: With each new prototype implemented in actual school settings, new challenges and solutions have been developed that have improved the quality, utility, feasibility, and efficiency of the *RISE* to meet school assessment needs.

A Pilot Intervention for Struggling Readers

Lowry Hemphill

Wheelock College

Objective: The *Strategic Adolescent Reading Intervention (SARI)* is a student curriculum and teacher professional development approach designed to raise literacy achievement for

eighth and ninth graders reading at a fourth grade level. The intervention also seeks to advance secondary teachers' ability to work with reading difficulties.

Perspectives: Reading success in middle school and beyond requires mastery of a wide range of literacy components including an adequate academic vocabulary and skills for learning new vocabulary through reading, familiarity with patterns of academic discourse, well-developed reading rate and the stamina to access longer texts, and competence at using a variety of comprehension strategies to make sense of sophisticated content. When adolescents have not acquired these literacy competencies, academic performance in all areas is at risk. An added perspective is that reading interventions for adolescents need to reflect the task complexity and domain specificity of literacy for academic purposes or the skills promoted will not transfer readily to success in content area classes. Administrations of the *RISE* diagnostic assessment revealed that many Boston middle and high school students had difficulty with basic aspects of reading. Despite a drive to improve graduation rates in Boston, few resources have been directed to older struggling readers. In SERP's surveys of Boston middle school teachers, most reported that they felt poorly prepared to teach reading. In addition, before implementation of the *RISE*, the district had not systematically identified adolescent with reading needs.

Method: The intervention was piloted in 5 Boston schools in 2008-2009. SERP collaborated with administrators and school leaders to screen 8th and 9th graders for reading difficulties and regroup those achieving at about a fourth grade level into smaller intervention classes. Intervention teachers participated in a semester graduate course in reading pedagogy and then implemented a semester-long intervention curriculum. Assessments of teacher knowledge, interviews with student participants, and gain scores on the *RISE* were used to document program effects.

Evidence/Results: Intervention teachers had initially weak understanding of basic processes involved in reading such as the components of reading fluency and the relationships between vocabulary knowledge and comprehension. Teacher understandings shifted substantially as a consequence of professional development and experience implementing the intervention. At the end of the course, students reported enjoying intervention components that involved partner and small group work, particularly when these included exchanging perspectives on texts with other students and led to a growing sense of mastery. Literacy gains for intervention students varied by teacher and setting (e.g. regular education, ELL) but were most evident for word reading, vocabulary, and basic comprehension.

Significance: Results suggest that the intervention holds promise for improving students' literacy competencies, but the professional development model needs strengthening to ensure gains across teachers and classrooms with different characteristics. Similar to *Word Generation*, school-level factors such as the principal's role in mobilizing support for the program also appeared to be important to the success of the intervention.