

**Getting Past Go:
The Use and Perceived Effects of
External Provider Packages in High Schools**

**Margaret Goertz, Matthew Goldwasser, Kate Hovde,¹ Jennifer Mueller, Matthew Riggan²
University of Pennsylvania**

Consortium for Policy Research in Education

**Paper presented at the annual meeting of the
American Educational Research Association, San Francisco, CA, April 10, 2006**

The research reported in this paper was conducted by the Consortium for Policy Research in Education (CPRE). Funding for this work was provided by the U.S. Department of Education's Institute of Education Sciences (Grant #R308A960003). Opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Institute of Education Sciences or the institutional partners of CPRE.

¹ All comments and questions regarding this paper can be directed to Kate Hovde at khovde@pobox.com.

² Authors are listed in alphabetical order.

Changing Times: How High Schools and External Improvement Organizations Work Together

PROJECT OVERVIEW

Each year more high schools are identified as underperforming due to failure to make adequate yearly progress. To raise achievement, new performance expectations require teachers, schools, and districts to move beyond aligning curriculum, reallocating time, and other conventional approaches. In response, they are relying more and more on external school improvement organizations. Although U.S. schools since the early 1990s have increasingly turned to external sources of assistance for improving academic achievement, literacy, graduation rates, and other key problems, the challenge of making changes in schools, and high school in particular, is well documented (Siskin, 2003).

To better understand this challenge, CPRE researchers have prepared four papers that examine the efforts of five external reform organizations and their interactions with high schools. We focus on provider design strategies and challenges, the uses and perceived effects of the reforms in high schools, the mutual impacts of communication networks and reforms in schools, and the nature and school leadership in the implementation of reform efforts. This project is particularly important and innovative because it makes the reforms and their impacts on schools simultaneous objects of investigation. Findings will allow schools and providers alike to select and strategize more carefully in order to maximize the potential for deep use of improvement strategies.

The research presented in the four papers draws from interview, survey, and observation data collected during 2004 and 2005 at 15 high schools across the country and from staff at five external assistance providers. The external reform organizations—High Schools That Work, First Things First, Ramp-Up to Adolescent Literacy, the Penn Literacy Network, and SchoolNet—were selected as representative of the types of external assistance found in high schools during previous CPRE research (see Gross & Goertz, 2005). The providers include two whole school reform models, two literacy programs, and one strategy to increase data-driven instruction. The 15 schools in our sample were selected based on recommendations from the reform organizations. Each provider identified three schools with which they had collaborated for one to five years. Ten of the study schools (two from each of the five providers) were in their first or second year of implementation. Five “mature” sites had worked with their respective provider for more than three years. In this way, we were able to examine several phases of a given reform.

Schools at the earlier stages of implementation were visited twice (at the beginning and end of the 2004–2005 school year) while the mature sites were visited once. During each visit, interviews were conducted with teaching and administrative staff at the school and district level. Staff members with both central and peripheral involvement with the reform were targeted. In total, our findings are based on approximately 380 semi-structured interviews lasting about 30–60 minutes each. During our site visits, we also conducted guided observations in a sample of classrooms. Interviews were also conducted with staff members at each of the reform organizations. These interviews focused on general questions related to the reform as well as

specific issues related to the schools in our sample. Interviews with fourteen provider staff members were conducted and analyzed for this project.

In addition, a survey was conducted with all teaching staff at each of the 15 sites. Using social network analysis, the survey provided information on the communication networks that exist within schools. The survey also contained items that allowed us to measure the depth and breadth of the implementation of a particular reform in each of the 15 schools. Our survey findings are based on 1,057 returned surveys, with individual school response rates ranging from 60 to 90 percent.

The interim findings of this project have been compiled into four papers (a fifth paper on the role of the central office will be forthcoming) that each shed important light on the design, use, and mutual impacts of external reform efforts and American high schools. Taken together, the papers provide a multi-faceted approach to describing the complexities of making change in high schools.

Citations

Gross, B., & Goertz, M. E. (Eds.). (2005). *Holding high hopes: How high schools respond to state accountability policies* (CPRE Research Report No. RR-056). Philadelphia, PA: University of Pennsylvania, Consortium for Policy Research in Education.

Siskin, L. (1994). *Realms of knowledge: Academic departments in secondary schools*. London: Falmer.

The four papers currently based on the data from this project include:

Goertz, M. E., Goldwasser, M., Hovde, K., Mueller, J., & Riggan, M. (2006). *Getting past go: The use and perceived effects of external provider packages in high schools*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Riggan, M., & Supovitz, J. A., & Hovde, K. (2006). *They come in all shapes and sizes: Leaders and high school reform efforts*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Shiffman, C. D., Massell, D., Goldwasser, M., & Anderson, J. (2006). *Design as intended, design as enacted: External assistance providers and high school reform*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Weinbaum, E. H., Supovitz, J. A., Gross, B., Cole, R. P., Weiss, M. J., & Ricalde, B. (2006). *Going with the flow: Communication and reform in high schools*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

1. Introduction

This paper presents interim findings on the use and perceived effects of five external provider reform models in 15 high schools. The high schools were nominated by the providers and range from early implementers to mature sites.³ We have chosen the word “use” over “implementation” advisedly: This paper is part of a quartet of papers which together paint a more thorough picture of the factors and dynamics affecting implementation from a number of different perspectives. Just as the Provider paper in this group of papers confined itself to the perspectives of the providers,⁴ this paper is bounded by the perspectives of school-level actors—principals, coaches, and teachers.

The paper is organized into four main sections: (a) our definition of use and our analysis methodology; (b) a largely descriptive section presenting findings by provider model on the use and perceived effects of reforms across sites; (c) a summary of our overall findings on use and perceived effects; and (d) a section discussing factors affecting use across all providers. This last section is by no means presented as the last word on factors explaining use overall; rather, it is a compilation of issues that emerged as important in analyzing our data within our admittedly limited perspective. Both the Social Network and Leadership papers in this study pick up some of these themes and broaden the overall understanding of implementation dynamics.⁵

2. Definition of Use, Effects, and Analysis Methodology

As a preliminary step to defining use, we reviewed a number of studies of implementation and specifically the implementation of externally provided reforms (e.g., Berends, Bodily, & Kirby, 2002a, 2002b; Corcoran, 2003; Datnow, Borman, & Stringfield, 2002; McLaughlin & Mitra, 2001). Like in the Provider paper, we found that Coburn’s (2003) redefinition of scale into four interrelated dimensions—spread, depth, ownership, and sustainability—provided a helpful theoretical departure point from which to examine use of reforms. As we began to delve into the conceptual underpinnings of these four dimensions and our data, however, we found it necessary to both modify and add elements. In particular, reform ownership in our data was both difficult to convincingly identify and tended to overlap substantially in discussions of depth and sustainability. We therefore chose to drop ownership as a separate dimension of use, placing it instead within depth. A discussion of how we chose to draw on the concepts related to reform use follows.

2.1. Spread

Our definition of spread is slightly narrower than Coburn’s (2003), which includes the idea of spreading “underlying beliefs, norms and principles to additional classrooms and schools” (p. 7).

³ After the provider abbreviation, the number 1 denotes a school that was in its first year of implementation, the number 2 a school that was in its second year of implementation, and the number 3 a school that was implementing for 3-5 years.

⁴ For the Provider paper, see Shiffman, Massell, Goldwasser, and Anderson (2006).

⁵ For the Social Network paper, see Weinbaum, Supovitz, Gross, Cole, Weiss, and Ricalde (2006). For the Leadership paper, see Riggan, Supovitz, and Hovde (2006).

Spread in this paper stays at a more superficial level and encompasses basic levels of exposure to the reform (i.e., who knows about it and who does not, who received training and who did not); procedural enactment of reform elements (e.g., formation of small learning communities, elimination of low-level courses); the existence or absence of monitoring and support mechanisms; and the issue of fidelity to “non-negotiables” and/or modifications of the model.

2.2. Depth

Our definition of depth begins with a basic level of awareness and understanding regarding the purpose and goals of the reforms. We borrowed this concept from McLaughlin and Mitra (2001), who argue that in order for externally-developed, theory-based change to begin to make inroads in a school setting, teachers must understand the first principles of the reform, the “why” as well as the “how.” Although our data stay at a fairly superficial level, the concept of “awareness” also incorporates an individual’s or school’s overall sense of the need for change. We also include in depth the value teachers and administrators ascribe to the reform. Thus teachers may understand and agree with the purposes of the reform, go through the motions of the reform, but disagree with the actual elements of the reform. Conversely, teachers may value a reform for reasons that have little to do with its stated goals.

At the next level of depth, we include changes in teacher or school-wide behaviors and the motivations behind them; for example, a teacher’s decision to seek out training, or changes in school communication patterns. Going deeper, we include the depth or extent of any changes in instructional practice.

Finally, at the deepest level, we include changes in teacher beliefs and school norms, and ownership of the reform. We note that the overall study methodology was not designed to thoroughly capture several of these elements, in particular changes in beliefs and ownership. However, we have retained them as conceptually important and present what evidence arose. We also note that there is probably mutual influence, or a feedback loop, existing among these different levels of depth, and that the stages are not necessarily sequential within a school. In other words, individuals can be at different stages at a given point in time.

2.3. Sustainability

Our theoretical definition of sustainability divided the concept into three tiers: (1) sustainability as manifest in the institutionalization of the reform in school structures, (2) in the language with which teachers and others talk about what they do, and (3) in the incorporation of core reform principles into all aspects of school and classroom organization. While a critical concept, sustainability presented a problem for us on a number of fronts. First, we saw sustainability, at least theoretically, as largely a product of the interaction between spread and depth, and thus there is inherently some overlap in discussion. Second, in practice, sustainability in education is rarely within the control of those most involved in a particular reform (i.e., teachers and principals). Finally, of our 15 sites, the vast majority had been working with the external reform organization for one or two years—hardly sufficient time to develop anything resembling sustainability. We therefore chose to retain the concept, but focus the discussion primarily on

sustainability perceived by school staff, rather than on the independent assessment we might make as analysts.

2.4. Perceived Effects

With regard to outcomes, or effects, the choice was made during study design to focus on effects of reform as perceived by school staff. We defined effects broadly, including perceived effects on behavior, morale, and teacher relations as well as a range of perceived effects on students. From a practical standpoint, however, we found considerable overlap between the discussions of changes in teacher behavior as part of depth and changes in teacher behavior as an outcome. To avoid redundancy, we chose to present perceived effects on teachers under depth, and to present separately perceived effects on students.

We also found useful Coburn's (2003) notion regarding the existence of interaction and mutual reinforcement between the dimensions of spread, depth, and sustainability. Drawing on other study findings (e.g., Berends, et al., 2002a, 2002b; Larson, 1999; Pascale & Sternin, 2005), we further hypothesize the existence of interaction or feedback between use as we defined it and perceived effects. A graphic representation of this framework is found in Figure 1 on the next page.

2.5. Analysis Methodology

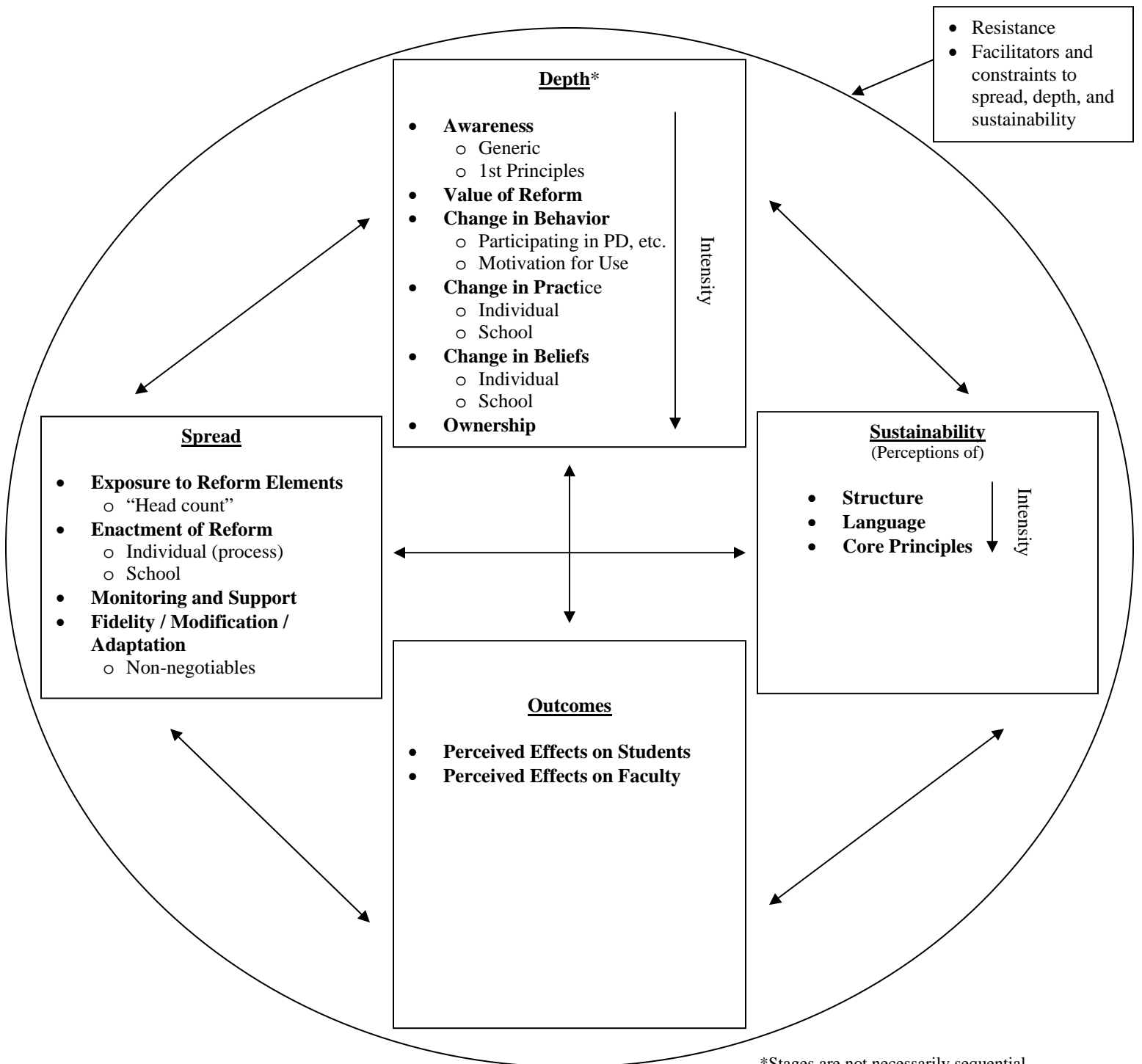
Our analysis drew on interviews with teachers and administrators as well as survey data from all 15 schools.⁶ Interviews were initially coded broadly in ATLAS.ti for discussion. References to “use and effects” were noted with the original Coburn (2003) definition in mind, but without any breakdown of concepts or a more precise attempt at definitions. We then looked at our use and effect data by individual school, with each team member taking responsibility for the three schools within a particular provider. It quickly became apparent that we needed a more developed conceptual framework to handle both the quantity and content of the data, and subsequent team discussions led to the development of the framework described above and presented in Figure 1 on the next page. Using this modified Coburn framework, we recoded our school-level data and wrote detailed case studies for each school. Survey data on aspects of use were also incorporated into the case studies for each school, and wherever possible helped triangulate findings.

With the school-level case studies in hand, we then did a cross-site analysis within each provider based on research questions tailored to our framework and school-level results. These cross-site analyses compiled and compared basic information on what did and did not happen in the schools, and considered both school- and provider-specific factors affecting use. These analyses were the basis of the very condensed descriptions of use and effect by provider that follow in the next section. We then looked at patterns among use, effects, obstacles, and facilitators across providers and identified overall conclusions. These conclusions are presented in Section 5 of the paper. With regard to factors affecting use, it should be noted that given the diversity of schools and providers, how these factors came into play was often very different: the discussion attempts

⁶ For an overall description of study design, please see the Project Overview. For a more detailed explanation of the survey methodology, see Weinbaum et al. (2006).

to note these differences while providing evidence for why we considered each factor important. Obviously with regard to a particular school, not all factors had equal explanatory weight; our purpose, however, was to illuminate commonalities.

Figure 1. Use and Effect Framework.



*Stages are not necessarily sequential. Individuals within a school can be at different stages at any given point in time.

3. Use and Effect by Provider

3.1. Ramp-Up to Literacy

Introduction

Ramp-Up to Literacy (RU) is a safety net program designed to bring 9th and 10th grade students whose reading falls about two levels below grade up to grade level. It consists of a teacher training program accompanied by a very structured and prescriptive curriculum, and is focused on the content and delivery of instruction in the classroom, although certain school-level adjustments need to be made in order for the program to be implemented according to the design. While stand-alone RU is not intended to become a school-wide program, America's Choice (AC), from which the program derived, is, and our results are muddled by the fact that our mature RU site (RU3) had the full AC program in the high school and our early RU site (RU1) was implementing AC at the junior high school, which is located within the same building and administrative structure as the high school.

The district played a role in introducing RU at all three sites and all have struggled with low student achievement. In the case of RU1, the district recommended both AC and RU to the school as a means of addressing low performance. In the case of RU2, the district mandated the use of RU and purchased the program for other schools as well. Although the district for RU2 did not purchase technical assistance from the provider, it did invest considerable time and energy in modification and support of the design. After the district-mandated breakup of a failing inner-city high school, the then-principal at RU3 recommended the AC high school program to the staff of the newly formed “school-within-a-school,” which voted to give it a try. AC was one of a small handful of programs being promoted and supported by the district at the time.

Spread

As would be expected from the model, in all cases but one, teachers responsible for teaching RU classes attended the required RU trainings, and began using RU in their classrooms.⁷ At all three sites, however, exposure to and engagement in RU ideas and techniques went beyond the small cadre of RU teachers. While the RU program itself was not well known at RU1, the principal was promoting the use of the AC class structure, which is similar to RU, in the high school across the content areas. RU2, which has been implementing the stand-alone RU for two years and which is a much bigger school, made RU a showcase program with a very wide reach. Although only eight teachers were teaching RU, everyone in the school knew about the program and one of the RU teachers together with the literacy coach led after-school professional development for content area teachers on RU structures and techniques. At RU3, all English teachers but one received RU training. The school also implemented the provider-developed math version of RU early on, of which the math department continued to use elements. The principal and literacy coach promoted the use of AC and RU structures and techniques across the

⁷ The one exception was a teacher who missed the provider's summer institute and instead received a three day make up training by the district.

content areas through a combination of internal and external professional development. All teachers at RU3 were familiar with the AC/RU structure and some of the AC/RU vocabulary and techniques.

With regard to fidelity to the model, RU1 and RU2 appeared to be observing its structural requirements. At RU3, there were significant structural deviations from the model, including a lack of targeting (i.e., all 9th and 10th graders were placed in RU despite some having lower or higher reading levels than the recommended two grades below), larger than recommended class sizes, compression or division of the block of instructional time, and abandonment of cross-site tutoring. Survey data from RU teachers on the frequency with which particular techniques were used indicates variance in instruction, despite the prescriptive nature of the program. At RU1, where only two RU teachers were trained at the same time, answers were fairly consistent. At RU2 there were two teachers who diverged significantly from the other six in terms of reported frequency; both were noted in the interview data as having problems with the model. In the case of RU2, the district also modified the RU design to include more writing and additional test preparation, as well as changing a number of literature selections. At RU3, there was less consistency in responses than the other two sites. Whether the greater variation reflects individual adaptation based on preference or perceived need, individual adaptation to the structural limitations to RU's use in the school, or changes in training over time is not possible to know from our data thus far.

With regard to the spread of RU ideas and techniques beyond the RU teachers themselves, the basic structure of AC/RU—the process of breaking up a class into a focusing activity, short teacher instruction, student work time (both individually and in groups), and a closure activity—was the most common element. Administrators at all three sites saw the structure as a means of shifting teachers away from the traditional lecture format and getting students more engaged. Discrete techniques and ideas such as “read-aloud, think aloud,” word walls, guided reading, independent reading, and use of rubrics were also being disseminated beyond RU teachers, but as a menu of options and not in the sequence specified in the RU curriculum.

Depth

RU teachers for the most part understood the objectives of the program and expressed a notable degree of commitment to it. All agreed that the program was sufficiently prescriptive, and that it inevitably led to some changes or alteration in instruction. Some teachers felt the program had dramatically altered their teaching, while others felt it merely rearranged many things they were doing already. Among the latter, virtually all felt the rearranging made sense, and that the structure and sequencing worked in the classroom. Survey data, however, indicate not only variance in the frequency with which teachers employed techniques, but also that those techniques that diverge farther from existing instructional practice (e.g., individual conferencing with students, or the teaching of decoding strategies) had lower rates of use than those that could be more easily incorporated (e.g., guided or individual reading).

Non-RU teachers received information about the design secondhand from administrators, coaches, or other teachers. Their understanding of program goals and specifics varied, as did their support for the program. At RU1, only 30% of survey respondents agreed that they

understood the purposes of RU, while only 7% agreed that RU provided them with useful ideas and resources about changing their classroom practice (66% were not sure).⁸ Teachers at this site were more familiar with the AC class structure and some techniques, and there was reportedly a good deal of resistance. At RU2, 82% of respondents replied that they understood the purposes of RU, and 66% agreed that RU provided useful ideas for changing classroom practice. At RU3, 76% of survey respondents answered that they understood the purpose of RU, while 44% felt that the program provided useful ideas and another 30% were not sure.

Changes in teacher behavior, expectations, and school structures were more evident at RU2 and RU3 than at RU1. At RU2, RU teachers had an overlapping lunch period and met with the literacy coach every day; time was also made for RU teachers to share their experiences with others. In contrast, at RU1 the two RU teachers worked largely in isolation. At RU2, the literacy coach credited the quality of the external training with helping to recruit strong teachers to teach the weakest students, rather than the reverse. She also felt the success of the program was chipping away at a culture of low expectations for students. At RU3, several teachers as well as the principal credited AC/RU with fostering increased discussion of instruction among teachers. Ownership of the program among RU teachers and coaches was also evident in a number of examples of volunteer efforts to disseminate the program more widely (at RU2), defend the program against scheduling changes (at RU3), and adapt the program either to address weak spots in the design (e.g., writing) or meet the needs of other groups (e.g., non or very low readers).

Sustainability

The perceived sustainability of RU was strongly linked with both leadership and test results in our data. At RU1, where AC/RU was still getting off the ground in the face of a good deal of opposition, many teachers thought that the program would be unlikely to be sustained if either: (a) the vice principal were to change, since she is the driving force behind it; or (b) test scores did not show improvement. At RU2, most teachers saw the program as a smashing success, and hence continuing. One respondent, however, worried about the district's commitment to the program and its ability to provide the same level and quality of training once the two-year contract with the provider expired. Here too, there was anxiety about seeing concrete results of the program in terms of test scores. At RU3, the principal's pursuit of grant money associated with another program was perceived by some of the faculty as being further evidence of a gradual abandonment of the RU program, if not RU ideas and techniques per se. Perceived sustainability at RU3 also differentiated between individual and institutional levels: Dedicated RU teachers claimed they would continue teaching the way RU taught them, no matter the level of institutional support. In fact, they had been in this position for some time, as the basic structural requirements of RU in the school had eroded.

Perceived Effects on Students

Across all three sites, teachers familiar with RU agreed the program motivated students to read, and that the structure of RU was very helpful in classroom management. Teachers at RU2 went

⁸ Survey respondents do include at least some of the RU teachers, who uniformly answered both questions affirmatively.

further, describing a culture shift in the school whereby it became acceptable for students to read. The effect of RU on classroom management was a huge selling point for the rest of the school at RU2. This experience contrasted with that at RU3, where the classroom management aspects were viewed as helpful, but student discipline remained a concern. Several teachers at both RU2 and RU3 also discussed how RU helped improve student self-confidence. RU teachers almost uniformly cited seeing positive results with students as a reason for continued use of the program. There was less agreement both within and between sites with regard to whether RU would ultimately improve student test scores or indeed whether the provider goal of mainstreaming RU students within two years could be met. Confidence was highest at RU2: All RU teachers and 70% of teachers surveyed school-wide thought the program would raise student achievement. Even at RU2, however, there was concern expressed about what should happen to students who complete two years of RU but are still not ready for a regular English class.

Facilitators, Obstacles, and Factors Affecting Use

Leadership, support, and the quality and detail of the provider training and materials were the most salient facilitators for the use of RU in all three sites. Few RU teachers expressed doubt about what to do, although there were occasional difficulties in actually doing it. Leadership and support took a number of forms, from ensuring the structural requirements of the program, to intensive instructional help from the coach (RU2, RU3) and district personnel (RU2). Provider support through the training program was cited by an RU teacher at RU1, while the provider also played an important role in the early implementation of RU at RU3 through the AC affiliation.

Policy and program churn at the district and school level was cited as an obstacle at all three sites: At RU1, a history of multiple programs that had come and gone fueled resistance to AC/RU; at RU2, a possible waning of district support for RU was seen as undermining sustainability; at RU3, the pursuit of funding through a different program was seen as “the next new thing.” Other obstacles cited tended to be more site-specific, and included scheduling problems at RU3 that undermined basic program requirements and a transitory student population with a high degree of absenteeism at RU2. The maturity of the program at a school was a more limited predictor of use than might be expected. For example, the structural components of the model were more faithful at RU2 than RU3, and frequency of use of techniques among RU teachers at RU2 appeared to be more consistent than at RU3.

3.2. Penn Literacy Network

Introduction

The Penn Literacy Network (PLN) is a non-prescriptive, literacy-focused program that targets individual teachers across content areas. PLN provides teacher training and support in literacy techniques through a number of voluntary courses, which normally meet weekly during the semester and carry graduate credit. Following the required introductory course (Course 1) teachers are able to sign up for subsequent PLN courses, many of which aim to help teachers better integrate literacy strategies into other academic disciplines. PLN prefers that teachers enroll in its courses voluntarily, and offers its courses at school sites to attract teachers. Also important to note is that because PLN is a non-prescriptive, teacher-focused program it looks

different in every school. Central to the three schools in our sample in particular is the aim of the reform: literacy training at the individual level (PLN1, PLN2) or whole school reform (PLN3).

PLN does not target schools, but gains access to schools and teachers through word-of-mouth recommendations by teachers or others. In the case of our three sites, each school had a point person—an individual responsible for bringing PLN to the school. The point person was not necessarily located at the school site. For example, at PLN1 this person was a district administrator and at PLN2 was an administrator at another school in the district. The role and level of engagement of these individuals varied by school.

Spread

The spread of PLN across the three schools was fairly consistent with the design's intent, focusing primarily on individual teachers. At the two early stage schools (PLN1, PLN2), all teachers received general information about PLN at an introductory presentation, however, only the small number of teachers at each school enrolled in Course 1 received continued training from PLN.⁹

At the mature school (PLN3), PLN took on a unique form. Due to changes external to the school (e.g., state and federal requirements) about five years ago, PLN3 was required to adopt a whole school reform program. At that time, the school had just started working with PLN in its traditional form, and successfully petitioned the state to have the design recognized as the school's whole school reform model. While the focus remained on individual teachers, using it as a whole school reform required modifications to PLN's roll-out strategy because all teachers were required to participate in PLN at some level. The design, therefore, evolved to include a range of professional development opportunities of varying intensity; for example, in-school professional development workshops, PLN courses, and teacher-led PLN sessions.¹⁰ As a result, engagement became two-tiered. One tier was comprised of teachers who were exposed to PLN at a superficial level (e.g., a list of PLN strategies distributed to teachers, PLN school-wide workshops). The other tier was made up of teachers who enrolled in PLN courses. Both Course 1 and Course 2 had been offered at PLN3, as well as three additional courses focusing on literacy, technology, and math. Moreover, engagement differed within these two broad tiers, especially within the second tier of teachers enrolled in PLN courses. Interview data show that teachers who took a course beyond Course 1 and/or Course 2 were more likely to enroll in more than one of the advanced classes.

With regard to fidelity to the model, enactment at PLN1 and PLN2 appeared to be consistent with PLN's traditional model. However, it was much more difficult to judge the level of fidelity at PLN3 because the model enacted is the only whole school version of PLN. It is unfair to imply that enactment at PLN3 was not faithful to the model when the intention was clearly for PLN to take on a different form at PLN3. What is unclear is the extent of the intended

⁹ Number of teachers enrolled in Course 1 at each early stage school: PLN1 = 7; PLN2 = 6.

¹⁰ Most of the professional development workshops were administered by PLN staff; however, teachers who were identified as experienced PLN implementers also assisted with the workshops. At least one teacher at PLN3 was a trained PLN facilitator, and a department head, who was an early adopter, held both informal and formal new teacher PLN workshops.

modifications and PLN3's fidelity to those changes. For example, monitoring is not a component of PLN's traditional model; however, it is available as an add-on to the traditional model for an additional cost. While PLN1 did not sign up for this option, PLN2 paid for this additional component called "mentoring" by PLN. This add-on included observations and feedback by a PLN staff member of teachers enrolled in Course 1 at PLN2. At PLN3, it was reported that PLN teachers were visited by a PLN staff member during the school day. Based on these statements, it appears that monitoring directly related to PLN occurred at PLN3. Yet, it is not clear whether PLN3 paid for monitoring (mentoring) or if this is a modification made as part of the whole school reform agreement between PLN and PLN3.

Depth

Awareness of the program goals and objectives reflected teachers' participation in the reform. When asked on the survey if they agreed with the statement "I understand the purpose of the Penn Literacy Network," 85% of teachers at PLN3 said they agreed or strongly agreed. At PLN1 and PLN2, the percentage of teachers school-wide who agreed or strongly agreed with the same statement was lower (47% and 58%, respectively). These results are not surprising because they reflect the different intents of PLN at each school. At PLN1 and PLN2 only a small number of teachers were enrolled in Course 1 and exposure for teachers school-wide was limited to an introductory presentation designed to recruit teachers for the course. Of participating teachers¹¹ at PLN1 and PLN2 enrolled in Course 1, 6 out of 6 teachers at PLN1 and 4 out of 6 of teachers at PLN2 said they either agreed or strongly agreed with the statement about understanding PLN's purpose. In interviews, participating teachers at all three schools had a relatively clear understanding of the goal of PLN: each gave statements consistent with PLN's stated focus on training teachers in literacy strategies.

In its more traditional form, PLN does not require any organizational or structural changes at the school level. Instead, individual teachers who take PLN classes are expected to exhibit changes in behavior and, most importantly, changes in instructional practice that reflect adoption of PLN's core principles and instructional strategies. Change at each school reflected the whole school reform nature of PLN at PLN3 versus the traditional, individual-targeted design of PLN at PLN1 and PLN2. Only the subset of teachers enrolled in Course 1 at PLN1 and PLN2 described changes in behavior, whereas most, if not all, PLN3 teachers reported changes in behavior and instructional practice as a result of PLN. Of the teachers at all three schools that reported changes, consistent themes emerged around generally being more literacy-focused and learning to think differently about how to structure and spend class time. One behavior change that did not seem to depend on the design of PLN at each school was teacher communication. Across the three schools, teachers said they communicated with teachers in other subject areas more as a result PLN.

Survey data show that participating teachers across the three sites felt confident with regard to their ability to use PLN strategies. However, teachers were split roughly down the middle about whether PLN required them to make major changes in their instructional practice. Interview and

¹¹ Participating teachers were teachers involved in PLN at each school. PLN1: 6 of the 7 teachers who took Course 1 and were interviewed also took the survey; PLN2: 6 teachers took Course 1 and were interviewed; PLN3: all teachers who were interviewed and responded to the survey.

survey data on specific PLN strategies, however, show that feeling capable and prepared did not necessarily translate into use of PLN. On the survey, the frequency with which PLN teachers reported using particular strategies indicates variance in actual instruction (which is not surprising given the non-prescriptive nature of PLN). Stated differently, despite feeling capable of making instructional changes and prepared to use PLN strategies in their classrooms, teachers incorporated the strategies at varying levels. Specifically, teachers tended to incorporate PLN strategies that aligned with their existing practice, while downplaying strategies that significantly challenged or altered it. PLN strategies with higher levels of use (e.g., pair/small group work, “do now”) did not require large departures from typical classroom practice; most were similar to or the same techniques already used by teachers. On the other hand, PLN strategies with lower rates of use (e.g., using primary source documents, note-taking technique for lectures) asked teachers to more significantly alter their teaching practice. The further a strategy required teachers to depart from their typical classroom practice, the less likely they were to use that strategy with regularity, irrespective of their sense of preparedness and capability.

Looking more specifically at PLN3, where all teachers were to be using PLN, the extent to which teachers changed their instructional practice to incorporate PLN strategies also appeared to be dependent on the type of PLN training a teacher underwent. In interviews, PLN3 teachers who attended only the workshops expressed difficulty executing the strategies, reported using only the ones they were comfortable with, and took fewer risks trying new strategies. Teachers who took at least one PLN course in addition to attending the workshops reported using more PLN strategies and were willing to take more risks trying new strategies in their classrooms. Essentially, different levels of exposure to PLN (workshops versus courses) had an impact on understanding, use, and consistency across teachers at PLN3.

Sustainability

At PLN1 and PLN2, voluntary participation in the design raised questions about sustainability. Teachers and administrators were unsure if additional PLN courses would be offered beyond the current school year. While literacy was a district-wide focus at both sites, PLN was seen as a professional development option for teachers to learn new literacy strategies. At the end of the school year, teachers at both schools speculated that Course 1 would be offered again the following year, with the hope of attracting more teachers so Course 2 could be offered the year after.

At PLN3, teachers generally attributed the “sticking power” of PLN to the fact that it is not very intrusive. Most teachers noted that of the available whole school reform models, PLN posed the least threat to their teaching practice and aligned well with the literacy goals of the school. In addition, most teachers felt that unless a remarkably better reform option came along, PLN would stay at PLN3. The principal commented that PLN was pretty much entrenched and saw no reason to change. It was noted by a few teachers that because the administration was pushing PLN, teachers would continue to use the strategies. However, one teacher said that in order to prevent PLN from “fizzling out” over time the school needed to focus on sustaining the reform through increased teacher support and continued training.

Perceived Effects on Students

Across all three schools, teachers credited PLN with having positive effects on students. At PLN1 and PLN2 teachers who took Course 1 spoke of effects on students in their individual classrooms, whereas at PLN3 teachers gave examples of effects on students both in class and school-wide. At the classroom level, teachers at the three schools said students were more engaged and motivated in class and were reading and writing better. Special education teachers highly praised PLN for its effectiveness with lower-performing students, and one teacher at PLN2 said she found the strategies to be successful with ESL students. At PLN1 and PLN2, teachers said it was too early to draw conclusions about overall changes in student and/or school performance. At PLN3, teachers and administrators credited school-wide changes in student performance to the use of PLN strategies by teachers. For example, increased scores on the state-mandated assessment and an overall greater emphasis on literacy in the school were attributed to the use of PLN.

Facilitators, Obstacles, and Factors Affecting Use

The perception by teachers that PLN is non-intrusive and easy to learn and use appeared to have facilitated its use by a majority of teachers at PLN3 and the teachers who took Course 1 at PLN1 and PLN2. However, this finding may be influenced by the fact that PLN's lack of prescriptiveness allowed teachers to selectively adopt strategies that did not challenge their existing practice.

Across the three schools, maturity was less of a factor than the intent of the reform. At PLN1 and PLN2, PLN was available as a literacy training and support opportunity for individual teachers—a professional development option—whereas at PLN3, PLN was a whole school reform required to involve teachers school-wide to some degree. At PLN3 in particular, training had a notable impact on teacher use of PLN. Teachers who took PLN courses not only used PLN strategies more often in their classrooms, but also used strategies that required a deeper level of understanding and commitment than teachers who only attended in-school PLN workshops.

Lastly, school resources (e.g., time) combined with the limited involvement of school leadership in support of PLN contributed to one of the greatest challenges to teacher use of PLN at PLN1 and PLN2: balancing content depth and curriculum pacing. Department chairs and/or administrators were generally not involved with PLN at these sites, and in the absence of guidance to help mitigate the tension between content depth through PLN strategies versus content coverage as dictated by curriculum pacing, teachers often ended up sacrificing the former for the latter. Moreover, the perceived need for change in instruction at these sites was low: most students at both schools were high achievers, and there was no internal or external mandate for reform.

3.3. First Things First

Introduction

First Things First (FTF) uses a prescriptive structural model (the creation of Small Learning Communities [SLCs] and a Family Advocate System [FAS]);¹² related professional development for teachers; and new leadership positions in the schools (School Improvement Facilitator (SIF) and SLC coordinators) to engender changes in relationships between and among teachers and students and in instructional practice. The model also relies on interactions within SLCs to spread and deepen the non-structural aspects of the reform, particularly the instructional component, Engagement, Alignment, and Rigor (EAR). The deeper values of the reform are intended to become apparent, and changes in teacher beliefs, pedagogy, and social relations are expected to occur, in the “doing” of the reform.

FTF-school partnerships were developed in different ways at each of the three schools. FTF recruited FTF3 to be in an early cohort of schools nationwide. In the case of FTF2, the district sought out a partnership with FTF to implement the reform in several schools; FTF replaced HSTW as a reform program at the school. At FTF1, a planning committee composed of school and district administrators selected FTF from a range of school reform options.

Spread

FTF requires schools to follow their roll-out script and timeline closely in the initial years of the reform. The program engages all school staff immediately by creating SLCs in the planning year, and implementing the SLCs and FAS groups in the first year of implementation.¹³ FTF works closely with school administrators to create SLCs, implement a process for student and staff selection into these SLCs, develop a new school schedule, and make the necessary physical moves to locate teachers within SLCs together. FTF also provides introductory training to all staff, and more intensive training to the SIF and SLC coordinators.

Although the three schools were at different stages of implementing FTF, each had created SLCs, assigned all teachers to SLCs and FAS groups, and begun to implement EAR in line with the design of FTF.¹⁴ SLC meetings had become a regular, even dominant, feature in the life of both FTF3 and FTF2. However, the substance of these meetings varied across the three schools reflecting, in part, the stage of FTF implementation at each school. While the SLCs at both FTF3 and FTF2 met twice a week, teachers at FTF3 reported that the meeting time focused primarily on student discipline, student engagement and student work, and on improving instruction. They spent less time talking about alignment, themes, and course content. Teachers at FTF2, which was in its first full year of implementation, reported that they talked about the full range of topics on a regular basis. In contrast, the SLCs at FTF1, which were in the planning stage, met only monthly and primarily discussed course content, curriculum alignment, and SLC themes. Interview respondents at both FTF3 and FTF2 reported that there was a range in the

¹² In FTF, within each SLC students are assigned to a teacher who becomes their family advocate working to ensure more continuity of care between home and school.

¹³ In our sample, FTF1 was in the planning year and FTF2 was in its first year of implementation.

¹⁴ At FTF1, FA and EAR were to be implemented in the 2005-2006 school year.

functioning of the SLCs across their schools. In FTF3, this was attributed in part to high teacher turnover and staff personalities.

The FAS had become part of the fabric of FTF3; teachers met weekly with students and meetings focused on what FTF expected (e.g., performance in class, progress toward graduation, and post-high school options). The FAS was not implemented consistently at FTF2. Teachers reported that time for FAS was pre-empted for SLC meetings or coverage of classes and that their contact with students was sporadic. Only 16% of the survey respondents met with all of their FAS students as a group at least once a week; half of the teachers met with students only once or twice a month. As discussed in the next section on depth, EAR was not well-implemented at either FTF3 or FTF2, the two schools in our sample enacting this component of the reform in 2004-2005.

With regard to fidelity, the three schools followed the FTF design closely, with two exceptions. First, although FTF calls for detracking within the SLCs, two schools made minor modifications to the FTF structure to address the needs of their students. To better serve their English Language Learners (ELL), FTF1 created a mini-SLC within their International SLC for 9th and 10th grade ELL students with core teachers who are ESL teachers, while FTF3 created a “sheltered transitional community” with the goal of moving students out of this institute by end of a school year into a thematic SLC with ESL support. At FTF1, a highly-tracked school, parental pressure pushed the school to schedule more courses outside the core of SLCs to enable students to take advanced placement, music, and other elective courses. FTF1 also plans to group honors students within FAS assignments so that their FAS period can be used for special seminars or electives. Second, FTF3, which was in its fifth year of implementation, did not have a SIF as called for by the design; instead the principal played this role.

Depth

Both survey and interview data show that teachers at all three schools felt they were aware of the goals of FTF and understood its purpose, although some teachers (about 25%) at FTF3 did not associate either the goals of the program or the program components with FTF specifically. This reflected, in part, the principal’s approach for communicating about FTF with his staff, which was to talk about the critical features of the program (e.g., FAS), rather than FTF per se. Although one goal of FTF is to increase the rigor of instruction and alignment of course content with state and local academic standards, most teachers across the three schools viewed the primary purpose of FTF as “keeping kids from falling through the cracks” because they are not academically motivated or engaged in their school work. Few teachers at any of the schools defined the goals of FTF as improving instruction. Teachers at FTF3 and FTF2, schools which were implementing EAR at the time of our site visit, viewed changes in course content and teacher pedagogy primarily in the context of getting students more engaged in their studies, rather than improving teaching in general. Across the sites, teachers involved in FTF leadership positions, who received direct training from the provider, were more knowledgeable about the multiple components of the program and more articulate about its goals, particularly regarding instruction.

Perhaps reflecting their understanding of the goal of FTF, teachers viewed the value of FTF primarily through the lens of relationships with and among students and their families. The SLC and FAS structures were generally viewed as appropriate mechanisms for building these relationships and creating conditions of learning that support students. Teachers reported that SLCs provided an opportunity for professional development and communication about both students and instruction. In addition, some teachers reported that SLCs led to closer personal relationships among teachers. Teachers also commented that this new structure weakened the structure of departments and limited communication among teachers in the same content area. Teachers credited FAS with enabling them to develop closer relationships with students, particularly at FTF3 where the program was well-developed and had a dedicated time for these meetings.

The reform appears to have had much less of an impact on instructional practice, perhaps because the instructional component of FTF—beyond its focus on student engagement—is relatively new (it was in the second year of operation at FTF3 and in its first at FTF2) and perhaps because it focuses on instructional techniques, rather than content. Teachers had mixed opinions of the value of EAR. At FTF3, only half of the teachers surveyed reported that FTF provided them with useful ideas and resources for changing their classroom practice or had exposed them to examples of student work the program seeks to foster. However, many teachers interviewed (many of whom were in FTF leadership positions) cited the introduction of EAR as a mechanism for teachers to reflect on and discuss their practice. At FTF2, as with the other two schools, there was not much focus on the instructional component of FTF, perhaps because of the large number of curricular initiatives coming from the district and the state. However, FTF2 teachers who responded to the survey were more positive about the potential value of the instructional component of FTF than those at FTF3.

The structural components of FTF have been institutionalized at FTF3 and, to a lesser extent, at FTF2. One measure of ownership of SLCs at both schools was that teachers felt too many students took classes outside their SLCs and wanted the SLCs to become more “pure,” that is, to increase the amount of instruction in the core areas students took inside their SLCs. There was much less, if any, ownership of the instructional component EAR, perhaps because it is a relatively new component of FTF and many teachers did not understand the purpose or use of it, and perhaps because teachers did not view it as adding much value to their current instructional practice.

Sustainability

Teacher and administrator perceptions of the sustainability of FTF in their schools were not linked in any way to a continued relationship with the provider. Rather, sustainability was viewed at FTF3 and FTF2 in the context of what was happening (or could happen) at the district level. A strong sense of ownership of and support for FTF by faculty at FTF3 did not, for example, necessarily translate into confidence that FTF would be sustained if the principal, who was a strong proponent of the program, was replaced. Even the principal noted that the continuation of FTF would depend on the philosophy of a potential replacement. Faculty support for continuing the reform at FTF2 appeared to be split, although there was support for keeping the SLC structure even if the program ended. The district and school administration at

FTF1 were strongly committed to implementing FTF and most teachers appeared willing to give the program a chance as everyone shared the conviction that the status quo was not working for a large portion of the student population.

Perceived Effects on Students

The staff at FTF3 felt students had built strong relationships with their teachers and administrative members of their SLCs, students had bonded with each other in their SLCs, parents knew they could call their child's teacher (and vice versa), and teachers had a better sense of students' needs. Although in place for only one year, some of the teachers interviewed at FTF2 spoke positively about the effects of FTF on students, such as improved student behavior or engagement and slight improvements in academic achievement. Teachers at neither FTF2 nor FTF3 could speak to student test scores because they did not have a full year's worth of data (FTF2), or changes in the state assessment made longitudinal change difficult to track (FTF3). Faculty at FTF3 also offered their perceptions of other student outcomes, including increased retention of students in school, increased graduation rates, increased numbers of students prepared for and going to college, an increased number of advanced placement courses to support students who want to go to college, and that students understood what is expected of them in class. These perceived effects on students at FTF3 reflect the intended goals of the provider and the expected goals of staff at FTF1 and FTF2. These goals underlie support for FTF among staff at the three schools.

Facilitators, Obstacles, and Factors Affecting Use

In spite of being at different stages of implementation and varying in their levels of human, social, and resource capacity, the three FTF schools rolled out the structural components of the design in the time frame and in the way intended by the provider. This uniformity reflects the prescriptiveness of the design, direct support from the provider, and strong advocacy of the program by the schools' administrators (FTF3, FTF2) or district personnel (FTF2, FTF1). Differences in the implementation of components of the model reflect lack of time (FAS at FTF2) and the extent to which the component challenged teachers' beliefs about instruction (EAR at all schools) or responsibility for addressing students' behavioral and social problems (FAS at FTF2).

3.4. High Schools That Work

Introduction

As a whole school reform design, High Schools That Work (HSTW) operates from the philosophy that change must occur from the bottom up. Schools are the design's primary target although its reputation has motivated whole districts to adopt its approach. HSTW is a non-prescriptive reform built upon 10 basic premises, including raising expectations for all students, better linking academic and vocational curricula and guidance/advising, and engaging students in challenging content. Although non-prescriptive, HSTW has two standard interventions all schools must accept: an on-site assessment (Technical Assistance Visit [TAV]) by HSTW and the formation of "focus teams" of all faculty to ascertain what changes need to be made and how

to proceed with them. Once HSTW staff make a TAV, the provider invites school staff to attend conferences and trainings and encourages the school to join a network of other HSTW schools. The provider also arranges for teachers to observe model HSTW classrooms around the country. Over time, HSTW has assembled a menu of research-based best practices from which schools may select techniques, resources, and technical assistance that best meet their needs. While on-site coaching and support is available to schools at an additional expense, HSTW training is largely a turn-key approach. That is, individuals who receive HSTW training off-site are expected to use what they learned to facilitate changes directed towards school-specific issues.

Ideally, schools decide to bring in HSTW on a voluntary basis. Usually a vote is held and at least two-thirds of the faculty must agree. This was not the perception of some of the teachers at HSTW1, who claimed they were unaware they were voting at the time, or at HSTW2 where some teachers felt like the decision had already been made in advance and their voting was strictly pro forma. In our sites, HSTW entered largely through relationships at the district level. At HSTW3, the assistant superintendent was the person most responsible for bringing in HSTW and maintaining its presence in the school. At HSTW2, the district wanted to shake up the direction of the school and recruited a new principal who agreed to come only if the school would adopt HSTW, a reform with which he had prior successful experience at his former school. After its first year, the district staff person most responsible for HSTW was appointed assistant principal at HSTW1.

Spread

After the initial TAV by HSTW staff, school faculty form focus teams whose task is to identify and strategize how to take action on identified needs (e.g., increase student engagement, better prepare students for post-secondary education and/or employment). At the three schools, all teachers either volunteered for or were appointed to a focus team. These teams developed different concentrations and plans for implementation. At HSTW1, focus team attention, training, and resources were applied to creating a guidance and advising program. At HSTW2, the teams created grade-level academies and worked to communicate across disciplines to increase rigor and raise curricular expectations for all students. At HSTW3, focus teams concentrated on five different areas of school function related to HSTW's core principles. After four years of participating with HSTW, the HSTW facilitator working with HSTW3 recommended consolidating to three focus teams to plan for student advising, professional development, and a holistic strategy to raise expectations and achievement for all students. Again, because HSTW can be used to engage in structural and/or instructional reform, and is premised on the belief that change best occurs when it is home-grown, the presence of HSTW enabled the schools to enact strategies that best met their needs as they collectively defined them with HSTW.

All three schools arrived at similar conclusions about their needs that corresponded with HSTW core principles. In an effort to “raise the bar” and increase academic rigor, all schools worked to eliminate their lowest level courses. HSTW1 and HSTW2 also standardized the syllabi within content areas and coordinated common examinations. To better engage students, each school experimented with QUILT, an instructional technique designed to create a more question-focused and student-centered approach to classroom teaching. To prepare students for post-high

school experiences, the schools experimented with improvement to their guidance and advising and used the creation teacher groupings (e.g., grade-level academies, departments, SLCs) to nurture and direct academic and vocational interests. Many high schools, not simply those involved with HSTW, wrestle with these same issues and employ similar strategies. Our data does not indicate whether the ideas and focus team concentrations came directly from HSTW or from elsewhere.

Because training is voluntary, the manner in which HSTW ideas are distributed and who benefits from them varies. Operating from the principle that change must emerge from the ground up, HSTW provides training to the faculty members who are the most receptive and open. Our data show no discernable pattern linking these individuals together, although all three schools made a point of sending their young and newest teachers to HSTW conferences. Also, it should be noted that there was a very limited amount of interview data from teachers on the details and content of their HSTW training, and our survey found that 50% of teachers said they had not received any training at all. In general, those who seemed slowest to respond and/or who had not taken advantage of training opportunities were the older and more seasoned teachers.

While schools did have the presence of HSTW facilitators on-site, who over time observed, coached, and made school-specific suggestions, their influence and role in oversight was not something teachers mentioned as significant. Use of HSTW was monitored largely by administrators and in-house leadership, the character of which depended upon the training and experience of these individuals. The principal at HSTW2 had considerable experience and exercised a heavy hand in terms of oversight. At HSTW3, two in-house staff developers employed by the district were available and active in their efforts to deepen the reform, but the assistant superintendent, whose office was housed inside the school, was the primary driver of the reform. At HSTW1, the facilitator for HSTW was an active and welcome presence, but even so teachers felt he reserved judgment and let them decide for themselves what to do when many felt that they could have had a speedier and more efficient adoption of the program if he had taken a more direct approach.

Apart from agreeing to the TAV and the formation of focus teams, the model has few specific components that schools must adhere to. Fidelity to the model, therefore, consists only in these two requirements and whatever actions are taken in pursuit of the 10 core principles. The model is designed to be open to adaptation and modification based upon what individual schools need. On one hand, this approach carried some ambiguity for teachers who felt impatient with HSTW's self-inquiry process and desired more leadership from HSTW staff. On the other hand, HSTW may have been attractive to some teachers for the same reasons; the reform did not tell them what to do and allowed them to define their own needs and solutions. The HSTW model does not specify many details and instead brokers a range of established best practices from which to choose. From a practical standpoint, because other reform initiatives were also present in the schools, it was sometimes difficult for teachers to ascertain whether a particular method or approach was recommended by HSTW or part of a bundle of things either chosen by the school or mandated by the district. For example, teachers at HSTW1 expressed confusion about whether the creation of SLCs and the use of a specialized literacy and vocabulary program came from HSTW or elsewhere.¹⁵

¹⁵ Our data indicated the SLC structure originated elsewhere.

Depth

Both our interview and survey data indicate that, almost without exception, teachers across all three sites had some awareness of HSTW. At a minimum, they knew they were in a HSTW school, whether or not they were fully on board. Because of the loose design structure, the actual depth of this particular reform was difficult to see. Likewise, while it can—and does—broker information about specific instructional practices, HSTW defers to the schools to ask for assistance rather than endorse or volunteer any preferred approach or technique. While all schools began with the formation of focus teams and problem definition, it was the work of each school's focus teams that largely determined the content of the reform and nature of the training sought among a HSTW-provided menu of best practices and training options. That said, most teachers readily concurred that HSTW is about “raising the bar,” increasing expectations for all students, and ensuring that all students are successful.

Our data are less conclusive about the extent to which any of HSTW's principles influenced the classroom-level behaviors and practices of teachers. For example, a few teachers were concerned that newly renamed courses, in an effort to eliminate lower level courses, still offered a lowered threshold of work and expectations in spite of the new course name (e.g., pre-algebra replacing basic math). Also, this commitment to a higher standard for all students conflicted with beliefs certain teachers had about students' abilities and openly raised doubts for others about the utility of holding all students to the same expectations.

A good example of depth was in the 9th grade academy at HSTW2. The 9th grade academy consisted of the most enthusiastic and experienced teachers, largely self-selected, who shared a perception that the school needed to do its utmost to prepare its students for success. Teachers in this academy were dedicated, extremely hard working, and fully committed to the principle that students will not fail, and they structured their academy and additional time outside of the school day to make that a reality. The experiences of these teachers contrasted sharply with those of their colleagues in the 10th grade academy who did not share the same commitment to the principle, were not a self-selected group, and were occasionally doubtful about the benefits of HSTW as a model for their work. Several of these teachers expressed doubt about the efforts going into providing additional support for students, arguing that in the “real world” failure is something that occurs. Certain students, several teachers reasoned, knew how the system worked and exploited it, performing when they felt like it, knowing that they would not be marked as failing. This was an unresolved issue exacerbated by the presence of HSTW.

Sustainability

Because HSTW is largely non-prescriptive, it appeals to schools on a philosophical level through its core principles. Several of these core principles—increased academic rigor, better student engagement in the learning process, and better student preparation for post-high school decisions—appeared to have staying power at the three schools. While actions to meet and uphold these core principles were expected to change over time, teachers felt these principles would be sustained. Also, teachers felt the focus group structure—a key vehicle for conducting needs assessment and decision-making—would be maintained. While there were some

dissenters, most teachers felt that HSTW was an asset to the school and hoped it would stay. Since the reform was only in its first and second year at HSTW1 and HSTW2, respectively, many teachers at those schools commented that, despite doubts, they wanted to give it a chance to take root and grow. Some said they appreciated a model that offered at least the promise of having some voice in decision-making, and knew that other reforms were not so open. A few teachers stated that what they learned at HSTW conferences was highly effective in their classrooms (e.g., questioning techniques) and they would continue to use the strategies whether or not HSTW remained. Finally, sustainability of the reform was bundled together with teachers' perceptions of the sustainability of school leadership. If the school leadership left prematurely, teachers thought that it might affect the presence of HSTW. Also, consistency of school leadership and leadership's commitment to HSTW as the school's reform model went a long way towards building buy-in from reluctant teachers and contributed to the perception of its overall sustainability.

Perceived Effects on Students

When asked about the effects of HSTW on students, the most noted effect was school-wide improvement on the state-mandated test. In addition, a few teachers mentioned that the number of "F" course grades decreased and the school's graduation rate went up. Reports of classroom-level effects were very limited; one teacher said he sensed a slight difference in student performance in his class, but felt changes would be more evident in a few years. This statement reflects HSTW's short tenure at the school (one year at the time of our site visit), and also raises questions about the strength of the relationship made by some between the reform and improved state test performance. Whether or not HSTW had any impact on test scores, the school, and more importantly the district, believed in the benefits of HSTW enough to look into expanding HSTW to other high schools in the district and contracting with the model's extension program, Middle Schools That Work, for all middle schools in the district.

Teachers at HSTW2 were more skeptical of the effects of HSTW on students. Teachers reported that while they did see some improvement in student performance (e.g., increased scores on the SAT and end-of-course exams), they were not sure they could attribute any changes to HSTW. A few teachers noted that the "newness" of HSTW at the school was likely why there was hesitation to link changes in student performance to the reform's presence at the school. At HSTW3, teachers generally thought HSTW had a positive impact on their students. Several technical teachers reported that their graduates seemed better prepared for their post-high school lives, but by and large evidence of specific impacts from HSTW remained thin. Interestingly, teachers overwhelmingly attributed improvements in student performance (e.g., increased test scores and improved overall student quality) to HSTW3 being a school of choice with entrance requirements.

Facilitators, Obstacles, and Factors Affecting Use

The location and style of school leadership had a strong impact on the presence of HSTW at all three schools. The HSTW2 principal and the two site coordinators at HSTW3 had a much more dynamic and enthusiastic relationship with their HSTW facilitators than did teachers at large. At HSTW2, the staff was heavily divided between those who were empowered and excited by their

principal and those who simply saw him as heavy-handed and inflexible. At HSTW1, leadership was still emerging and had not really bubbled up from the bottom as the HSTW model prefers. There were also examples of staff resistance to the reform. These included teachers who defended their status and did not see any reason to change, or shoehorned specific pieces of HSTW into claims that they were already doing it in their classroom(s), and those who were doubtful of the reform itself and were slow to accept it. A number of teachers, especially at HSTW1, expressed concern and surprise regarding how labor intensive it was to implement HSTW. It reportedly took a great deal of time to attend all of the planning meetings as well as the continued preparation of class and subject materials, documentation of efforts, and review of research. In contrast, at HSTW3, teachers did not characterize time commitment as a particular obstacle. It may be that as schools get deeper and go longer into the reform, time becomes less of a problem.

3.5. SchoolNet

Introduction

SchoolNet (SN) aims to build school district capacity to improve curriculum and instruction through the use of technology. Specifically, it offers several web-based applications designed to improve teachers' ability to use data to plan instruction, and align standards, curriculum, and assessment. These products include Account, a student performance database; Align, a tool to link locally developed curriculum with state or district standards; and Assess, a product designed to measure student progress toward meeting local standards. The district is the primary point of contact for the provider, and is the only recipient of direct services. SN does not specify any implementation process at the school level. The underlying assumption in this approach is if district staff are properly trained in the use of SN products, they will be able to pass that expertise along to school staff, who in turn will pass it along to teachers. In all three of our sites, the district selected SN to work with multiple schools. While many teachers were unfamiliar with SN, the most common perception among school staff was that it had been brought in by the district to improve access to and use of student performance data. Some associated this with compliance with No Child Left Behind (NCLB) legislation. The schools did not have input into the decision to contract with the provider.

Spread

The three districts included in the sample used a similar strategy to introduce SN to the schools. Each identified a team of teachers, department chairs, and/or administrators from several schools in the district, then provided them with SN training so that they would in turn train teachers in their respective schools. In all three sites, the team designated to become school-based trainers did not themselves receive sufficient training to carry out this task—most recipients suggested they did not know the applications well enough themselves to effectively train and support others. As a result, teachers received infrequent and/or superficial professional development around SN.

Survey data suggest low levels of engagement with the design at all three sites. At SN3, less than 23% of survey respondents described themselves as “involved” or “very involved” in SN.

At SN2 and SN1, these figures were 36% and 42%, respectively. Likewise, use of SN by teachers was limited in all three sites. At SN3, 87% of survey respondents reported that they had logged into their SN account “never” (70%) or “a few times a year” (17%), compared to 68% at SN2 and 56% at SN1. Interestingly, this pattern indicates lower levels of use in sites where the design has been the longest.

With regard to fidelity, interview data reveal a significant modification of the design in two sites. At SN2 and SN1 (the two sites at which there was some evidence that SN was being used), teachers were given data generated through SN, but were not actually accessing the program themselves.

Depth

Awareness and understanding of SN varied significantly across sites. At SN2, 76% of survey respondents agreed that they understood “the purpose of SchoolNet;” at SN3 and SN1, about half of the teachers reported that they understood the purpose of the design. Interview data are relatively consistent with these findings. In most cases, teachers were able to offer some explanation of what SN was and what it was intended to accomplish. Their explanations, however, varied significantly, and almost none seemed to have a complete understanding of all that SN offered, or how the various products related to one another. Most teachers’ understanding of the design was based on their awareness of Account, SN’s program for storing and organizing student performance data.

Survey data suggest that at all three sites there was considerable school-level support for the idea of data-driven instruction. Teachers believed data was an important tool for improving their instruction and better supporting students, and generally felt they had the technical skills needed to use data effectively in their classrooms. However, this support did not necessarily translate into support for SN, or recognition that SN might help schools become more data-driven. Similarly, survey data suggest that teachers at all three SN sites were at least somewhat supportive of specific data-driven instructional practices, such as using data to identify students in need of remediation, tailor instruction to meet individual student needs, and/or identify areas in which they could strengthen their content knowledge or practice. There is little or no evidence, however, that the adoption of SN increased the use of such practices.

Individuals who received the most training and exposure to SN seemed most optimistic about its potential contribution to the school. They suggested that access to students’ academic histories helped teachers to better understand students’ strengths and weaknesses. Compliance was another area in which SN was valued. Some school staff seemed to view data-driven reform as an external mandate reinforced by NCLB, and viewed SN as a tool for helping them comply with the law.

There is more evidence that SN supported changes in the practice of department chairs, assistant principals, or others charged with accessing, organizing, and distributing data to teachers. In these cases it appears that improved access helped them become more focused on data, and better able to share data with teachers.

There was very little ownership of SN at any of the sites. At SN3 there was overt animosity toward the design, largely fueled by poor relations between the school and the district (which contracted with the provider). At the other two sites the lack of ownership had more to do with teachers' limited exposure to the design and the resulting lack of awareness of the capabilities of SN with respect to data-driven reform.

Sustainability

At SN3, teachers indicated that SN did not enjoy broad-based support among the faculty; it seems clear that if put to a vote, the design would likely be eliminated. At the other two sites the results were less predictable. There was a general lack of awareness or understanding of the design among most teachers, but there was also significant support for the idea of data-driven reform practices.

Perceived Effects on Students

There is little evidence that SN in any way influenced teacher perceptions of student behavior or performance. Asked whether they agreed that “the changes called for by SN are helping or will help my students reach higher levels of achievement,” 34% of survey respondents at SN1 reported that they did not know, 27% disagreed, and just 19% agreed. At SN2, 39% of respondents reported that they did not know whether SN would help their students, 30% disagreed, and 24% agreed. At SN3, 43% of respondents did not know, 35% disagreed, and 19% agreed. There is almost no evidence in the interview data that SN had influenced perceptions of student outcomes.

Facilitators, Obstacles, and Factors Affecting Use

Three types of obstacles undermined the use of SN. The most common problem was technology access. At SN2, the majority of teachers did not have access to the internet in their classrooms, while at SN3 teachers struggled with old computers that broke frequently, were slow to be fixed, and ran slowly when they were operational. Because SN is based on teachers accessing online resources from their classrooms, this lack of infrastructure seriously undermined both the spread and depth of use. SN1 did not report any school-level technology problems.

The second most commonly cited obstacle was that the software was not user friendly. Teachers at SN3 reported it was time consuming to access student data, and it took too long to enter lesson plans into the online database. At SN2, teachers commented that the software was difficult to figure out, even for teachers who reported that they were technologically proficient.

A third obstacle was integration of data into SN. At SN2, student performance data were not updated in real time, making it more difficult for department chairs to make effective use of SN reports. At SN1, there were incompatibility issues between district data generated through a different assessment system and the SN database. At SN3, student data were updated infrequently; teachers complained that on the rare occasions that they did look at student records (primarily during training sessions), they were out of date.

Finally, school-district relations played a crucial role in the use of SN. At SN2, for example, a close, personal relationship between a district staff person and a school administrator helped build school-level support for SN. At SN3, a history of hostile relations between the high school and the central office significantly hindered the implementation of all district initiatives at the high school level.

4. Major Findings

This section presents major findings on the use and effects of the five reform models across our 15 sites and identifies factors that appear to explain differences in use both within and across the models. Section 5 of the paper discusses these explanatory variables in greater detail.

Fourteen of the 15 schools were using the reform models, but the extent of use and fidelity varied both within and across models.

With the exception of SN3, all of the schools were using elements of their reform models. All three FTF schools had created SLCs and Family Advocate System (FAS) groups and were teaching and advising students within these structures. All of the HSTW schools assigned faculty to focus teams and adopted, at the recommendation of these teams, structural (e.g., grade level academies) and/or curricular (e.g., eliminating lower level courses) changes. The three districts that adopted SN provided some training to administrators, and some teachers in SN1 and SN2 reported having used a component of the program. Teachers responsible for teaching RU classes at their schools attended the requisite trainings and were implementing RU in their classrooms. Teachers who took PLN courses similarly reported applying some of the techniques in their classes. However, at RU and PLN schools, teachers were less likely to adopt instructional strategies that required large departures from their typical classroom practice. Not surprisingly, given differences in model designs, we found greater spread in schools using whole school reform models (FTF, HSTW), and more depth in those using models that target a small number of teachers (PLN, RU).

The extent of use varied, however, across schools within a model, as well as across models. Use of SN, for example, varied across the three schools (from some to none), reflecting differential access to training and technology, as well as faculty support for the program. Across all sites, the intensity of use was related to the level and quality of training received by teachers. We also saw adaptation of the models in the context of our schools, some condoned by the provider and some not. In our sample of schools, model adaptations made in conjunction with a provider were in response to specific district and/or school needs. The relationship forged between PLN and PLN3 is a clear example of a school and its district working with a provider to adapt the model to fit the school's needs and resulted in a higher level of use. As noted in the PLN case study (Section 3.2), PLN3 was required to adopt a whole school reform model by the state and successfully petitioned to have PLN recognized as such with some modifications. FTF1 worked with the provider to place more courses outside SLCs to accommodate parental pressure for fine arts and advanced placement courses. In the case of RU2, the district made modifications to the model to incorporate state writing standards and made some changes to the RU literature

selections. While the provider was clearly aware of some of these modifications at RU2, it is not clear from our data to what extent these modifications were approved.

Model adaptations present at some schools were *not* the result of a joint effort between school and provider, but instead resulted from decisions made at the school and/or classroom level. At the school level, RU3 was not faithful to the structural elements of RU; for example, class sizes were larger and included students above and below the called for two grade levels below in reading, and block time was compressed. The RU design also underwent changes at RU2 at the school and classroom level, where RU became a showcase program and teachers school-wide were encouraged to use some of the design's structural and instructional techniques. Budget cuts led to a reduction in the number of SLCs at FTF3 as well as limited the "purity" of SLCs at FTF3 and FTF2. FAS groups at FTF2 also met less frequently because of scheduling problems. In our data, classroom-level adaptations were more detectable at the schools implementing instructional models (RU, PLN). Although RU is a prescriptive model with a fairly scripted curriculum, self-reported survey and interview data show that teachers were not always able to carry out the prescribed lesson plans as laid out by RU. Classroom management, incomplete resources (i.e., leveled books), among others, were reasons cited by teachers for classroom-level modifications. And at the three PLN schools, classroom-level adaptations were often due to, for example, uncertainty about using a particular PLN strategy and/or curriculum pacing demands.

Spread was easier to achieve than depth, but in some schools depth was accompanied by a spread of reform ideas.

FTF and HSTW engaged all teachers at a school in some aspect of the reform program, either through structural reorganization (e.g., assignment of all staff to SLCs in FTF) or mandated planning activities (e.g., focus teams in HSTW). As a result, most if not all teachers in these schools were exposed to the reform elements, generally aware of the reform, and participated in some aspect of the design. However, while these organizational changes are designed to establish a supportive environment for personal growth and instructional change, and some teachers acknowledged that the new structures provided an opportunity to discuss student work and their own practice, we did not see much evidence of instructional change at our sites. This could be attributed to a number of factors, including the lack of a direct focus on instruction, insufficient training in new instructional practices, and teachers' ambiguity about the need to change their practice to implement the reform. That is, the reforms did not sufficiently address teachers' zones of enactment (Spillane, 1999), a concept discussed in greater length in Section 5.4.

We did not expect to see much spread of PLN and RU, two models that target a small number of teachers at a school. At PLN3 and all three RU sites, however, exposure to and engagement in PLN and RU ideas and techniques extended beyond a small core. At PLN3, the design was implemented school-wide in response to a state mandate that the school adopt a whole school reform model. In all three RU schools, administrators promoted the use of RU structures and techniques across their staff as a way of moving teachers away from traditional lectures, and they facilitated teacher training in RU techniques. It should be noted, however, that in two of the RU sites (RU1, RU3) the additional presence of America's Choice contributed to spread.

Sustainability of the models was in question at all sites.

The sustainability of the reform models was questionable at all of our schools, including the mature sites. In some cases, sustainability was hindered by a lack of institutionalization of the program at the school. This was particularly true of SN. The program was used by few faculty at SN3, and poor relations between the school and the district fed animosity toward the design. While teachers at SN1 and SN2 supported the concept of data-driven decision-making, they were generally unaware of SN's capacity to address their data needs. PLN1 and PLN2, at which teachers viewed PLN as only one selection on a menu of professional development options, chose not to offer courses in 2005-2006, despite each district's commitment to teaching literacy.

In other instances, despite administrator and teacher support for the program, external factors, such as a reduction in funding or administrative turnover, threatened continuation of the reform. Teachers at FTF2 and FTF3, for example, wanted to preserve SLCs, which they felt permitted a greater knowledge of students and the creation of a more caring climate. However, teachers were uncertain whether the structure would survive a change in district priorities or school leadership. Similarly, teachers at RU2 and RU3 spoke about the positive effects of the program on students as a reason for continuing RU, but were concerned whether the program would stay when grant money expired. Reductions in state aid led PLN3's district to reduce the amount of PLN-related support it could fund, including professional development, mentoring assistance from the provider, and direct support to teachers from district coaches and facilitators. However, many teachers involved in instructional reforms (PLN, RU) stated that they would continue to use the instructional techniques they learned from these programs no matter what happened to the reform at the school level.

Many teachers felt the reforms helped their students, but recognized that improvements in test scores are the ultimate measure of any reform.

Many of the interviewed teachers at schools working with four of the models (FTF, HSTW, PLN, and RU) identified positive effects of the reforms on their students. These effects included greater student engagement and motivation in class, stronger relationships (FTF), improved reading and writing skills (PLN, RU), and increased retention in school (FTF, HSTW). Teachers who adopted or used a reform model did so in part because they perceived the model to be having desired effects on their students. In our data we found considerable evidence in support of a feedback loop between use and perceived effects. However, acceptance of the reforms by school staff and, ultimately, the sustainability of the programs require concrete data. Accountability pressures were evident at the vast majority of schools: Teachers and administrators were keenly aware that ultimately the "value" of any program was going to be judged with regard to increases in test scores. Increases in test scores during the tenure of a program were quickly seized upon as a reason for continued or greater use; a lack of improvement in test scores was something that dedicated "users" of a program needed to explain away and/or gave rise to efforts to adapt the program (e.g., efforts at RU3 to include more writing).

Maturity did not matter the most.

The maturity of a reform (i.e., its length of time) within a school would seem to be a logical explanation for higher levels of use across the designs. However, in our data, maturity did not bear out as strongly as we expected. In only two of the five designs did the mature site (PLN3, RU3) show evidence across all aspects of depth as we have defined them. While in all designs, except SN, the more mature schools displayed signs of greater spread of participation in the design elements, the RU and HSTW schools in their second year of implementation and the first and second year FTF schools, also achieved spread as defined by the design. Thus, although maturity is an important factor in the life of a given reform, other variables play critical roles. These variables include (1) the design emphasis; (2) training; (3) leadership; (4) zones of enactment; (5) school resources; and (6) a sense of urgency, and are discussed in detail in the next section.

5. Explanatory Variables

5.1. Design Emphasis

Each design employs a theory of change calling for specific activities to be undertaken, assuming that certain outcomes or processes will unfold from those activities and ultimately advance the goals of the reform. In so doing, each design stresses certain aspects of school organization or practice while deemphasizing others. This emphasis is evident in the way provider resources (e.g., staff time, materials) are developed and deployed, and in the changes in school organization or instructional practice required by the designs.¹⁶ From the perspective of schools implementing these reforms, design emphasis is most evident in the degree to which certain elements of each design are prescriptive. While the overall level of prescriptiveness in the designs varies considerably, most require certain steps be taken in the implementation process. These prescribed components or processes are, in effect, the reforms' levers for change.

FTF and HSTW are both whole school reform models stressing shared decision-making and increased student engagement. Yet the manner in which design components or processes are prescribed indicates that they approach organizational change in different ways. FTF is highly structured, establishing detailed, non-negotiable timelines for the creation of SLCs and a Family Advocate System (FAS), and for the introduction of instructional improvement strategies (EAR). While they participate in the identification of themes for SLCs, teachers otherwise have little input in the process itself. The logic of this approach, as explicitly stated by the provider, is that teacher buy-in comes from perceived benefit or success, and that such success will only be evident once the structural components of the design are in place (Connell, 2002). The HSTW design is more open-ended and process-oriented, but still begins with two prescribed activities: the completion of a Technical Assistance Visit (TAV) and the formation of focus teams. Both of these activities are designed to spark teacher dialogue and debate about the most important problems facing the school, and to build consensus around an agenda for change. This agenda dictates which structural aspects of reform should be adopted. In sum, HSTW emphasizes

¹⁶ For a more discussion see Shiffman et al. (2006).

teacher buy-in assuming it will lead to structural change, while FTF emphasizes structural change assuming it will lead to teacher buy-in.

Consistent with existing research on school reform implementation (e.g., Berends et al., 2002a, 2002b), our analysis suggests that components of the designs that were emphasized through prescriptiveness were most likely to be used by schools. All three of our FTF schools had the design's prescribed components (SLCs and FAS groups) in place within the appropriate time frame, and to a lesser extent teachers were using SLC and FAS time as expected by the provider. Similarly, at all three RU sites, teachers were using the curriculum, had adopted the instructional format required by the design, and were scheduled into longer class periods.¹⁷ In the HSTW sites, all teachers had participated in focus team meetings, and all three schools had completed their TAV from the provider. As a design, PLN is non-prescriptive. Teachers volunteer to take PLN courses, and use PLN techniques at their discretion.¹⁸ While the roll-out of SN at the district level is more structured, it includes no prescribed activities at the school level.

While it appears that prescriptiveness leads to use, our analysis suggests that prescribed components do not necessarily precipitate the kinds of changes intended by the design. In the case of FTF, we generally found that formation of SLCs and FAS groups led to the development of closer relationships among faculty and staff; in this respect the logic of FTF's emphasis seems to be supported in our data. On the other hand, SLCs are also to be sites for professional development related to FTF's instructional component (EAR), and teachers are expected to change their instructional practice by providing and receiving feedback and sharing strategies and experiences with their colleagues. There was little evidence in our data that this type of professional learning occurred, or that teachers changed their practice as a result.

With respect to HSTW, the TAV and formation of focus teams did not always have the intended effect of galvanizing the school community around an agenda for change. At HSTW1, while teachers acknowledged that the formation of their focus teams contributed to building new cross-disciplinary relationships they did not contribute (yet) to cross-disciplinary teaching or any major school-wide initiative. Teachers felt a little deceived that the promise by administrators of no additional work associated with HSTW was anything but true. Further, a number of teachers felt the problems facing HSTW1 could be more easily resolved via administrative fiat and certainly better administrative follow-through than the time spent meeting in focus teams, which instead could be spent with students and in their own classrooms. And while there was clearly a mandate for change at HSTW2, it came more from the principal than from the HSTW process. His leadership style, as much or more so than his commitment to HSTW, polarized many faculty members.

Because its curriculum and instructional practices are highly specified, RU did lead to changes in instructional practice for teachers using the design. Even here, however, the adoption of prescribed components did not necessarily have the intended effect. In some cases the design was adopted mechanically—teachers seemed to be “doing” RU without necessarily “getting” it.

¹⁷ At RU3, as noted, there were scheduling problems, and class periods were not as long as required by the design.

¹⁸ In the case of PLN3, all teachers were required to participate in PLN professional development (offered either by school staff or by the provider). This requirement was not imposed by the provider, however, but rather by the school and district to meet state-mandated whole school reform requirements.

This was particularly evident in the manner in which specific components of the design were adopted and used. As noted below in Section 5.4, elements that conformed to existing teacher practice appeared to be used more than those that significantly altered or challenged it. Yet it is precisely those more challenging practices, designed to help teachers diagnose and solve reading problems on an individual level, that are at the heart of instructional change intended by the RU design.

Instances of disconnect between adoption of prescribed components and the resulting changes (or lack thereof) intended by the designs raise two important points with respect to factors affecting use. First, it seems that prescriptiveness is a double-edged sword. On the one hand, it seems to ensure a measure of compliance in adoption. On the other hand, that compliance sometimes looks like “going through the motions” rather than making substantive changes in school organization or instructional practice.

Second, the relationship between adoption of specific design components and the resulting outcomes is grounded in provider assumptions about school capacity, defined as both the “will and skill” of teachers or administrators and the material resources available to them. With respect to the latter, the designs were significantly undermined when those resources were unavailable (further discussed in Section 5.5 on school resources). With regard to the former, the designs make different assumptions about both the skills possessed by school staff and the will to change school organization and/or instructional practice. For instance, RU requires schools to create complex schedules that allow for students to be grouped in specific ways with particular teachers, and SN assumes that teachers have the skills to use computers and the internet. There also appears to be considerable difference in assumptions about the instructional knowledge of teachers. FTF and HSTW aim to establish structures and processes whereby teachers can share instructional expertise; in the case of HSTW this expertise may come from outside the school if so decided by the teachers. The assumption appears to be that as a group, teachers possess the required instructional expertise to identify weaknesses in teaching and take appropriate steps to address them. Conversely, PLN and RU appear to be based on the assumption that teachers need specific guidance on instructional practice. In these designs, instructional expertise is not a latent quality to be surfaced from within the school community, but rather requires an external stimulus for change. Finally, all of the designs are based on the fundamental assumption that at some point in the implementation process, school staff will recognize the need to change and come together to support reform. Yet, as discussed below in Section 5.6, such will to change was not always evident. In its absence, use of the designs tended to remain superficial, even when spread throughout a school.

5.2. Training

It almost goes without saying that whenever a new reform model is introduced into a school a certain amount of training and technical assistance is essential to ensure that it is accurately and confidently enacted, and all of our providers make training a key dimension of their work. The costs associated with training and technical assistance affect what providers offer and/or schools receive, as well as what strategies and targets are chosen. Our data confirm many other studies’ findings that the quality, quantity, and depth of the training and technical assistance offered as well as the training strategy affects the degree to which schools and/or districts can push the

reform deeper and/or wider across their school environments (e.g., McLaughlin and Mitra, 2001; Corcoran, 2003).

Direct versus Turn-Key

The provider models differ with regard to who is trained directly by the provider, versus who is trained by others (train-the-trainer or turn-key model); several providers also rely on both types of training. Although a turn-key model is far less costly and may facilitate greater spread of reform, we found a clear relationship between direct training by the provider and greater depth of use. At FTF schools, for example, there was a much greater degree of understanding of the reform goals and stages among SLC coordinators trained by FTF than among teachers at large, who received information and training largely from the SLC coordinators. Teachers who received direct training through RU or PLN courses also exhibited a deeper understanding of and commitment to the reform and reported more instructional changes than those who had been exposed to workshop versions of PLN (in the case of PLN3) or modified versions of RU (occurring at RU2, or America's Choice / RU occurring at RU1 and RU3). HSTW relies on teachers attending summer institutes or other HSTW sites to come back and share information gained with colleagues. Our data of HSTW teachers' perspectives on and use of training is thin, but it did not appear that this was universally happening. In its standard form, SN trains district staff in the use of its technology, relying on them to train school-level administrators and staff. At all three sites, school-level staff did not feel themselves sufficiently well oriented to conduct the training at the school level, and teachers did not find the trainings particularly helpful.

Voluntary versus Mandatory (Incentive Structures)

Another element influencing the effect of training on use was whether the training offered was voluntary or mandatory, and what incentives existed for participation. Here too, mandatory training contributed to spread, while some degree of voluntarism seemed to contribute to depth, but could impede spread. Among our providers, FTF relied most heavily on mandatory training of the entire school staff. Although the turn-key training strategy created some inequities in information, the mandatory nature of the training put everyone on a more equal footing and helped ensure compliance with at least the structural aspects of the reform. In contrast, HSTW (the other whole school reform) espoused a voluntary approach to training, which enabled those teachers not particularly interested in or enamored with the reform to avoid deeper exposure. At the two early PLN sites, the voluntary nature of the training and lack of additional push for the reform from the leadership structure led many teachers to regard the training as merely a professional development option among several others. The fact that PLN offers graduate credit for courses also had a strong influence on who volunteered: several teachers confessed to signing up for the course because they needed the credits, and being surprised that the course was as good as it was. At PLN3, the discovery that the "credits" did not fully count towards a master's degree fueled a backlash against taking the semester long intensive courses. It is also noteworthy that at PLN3, some exposure to PLN through in-school professional development workshops was in fact mandatory to meet state-mandated whole school reform requirements.

For SN and RU the issue of mandatory versus voluntary was less clear, since under the SN design that decision is left up to the district and RU targets only 9th and 10th grade English

teachers. There was clear resentment among teachers at SN3 that what the district had advertised as a voluntary training later turned out to be mandatory. With regard to RU, there was no evidence in our data of a teacher being forced to take the training, but there was evidence of principals and coaches persuading teachers to do so. At RU2, the attraction of high quality external training where teachers were treated like professionals was used as a tool to recruit strong teachers to undertake the program.

Ongoing Support / Technical Assistance

A third element related to training and affecting use was the existence, frequency, and character of ongoing support and technical assistance at the school and classroom levels. SN, for example, had no intervention at this level. At the FTF and HSTW sites, provider staff visited the school on a fairly regular basis to provide training, support, and technical assistance, particularly during the early stages. Yet in concert with the models, the character of the technical assistance was quite different and frequently site specific. For example, at FTF1 the provider responded to school requests for specific help on at least one occasion, but the focus was on how to complete the predetermined tasks laid out by the model. The HSTW site facilitator also made fairly frequent visits to HSTW1, but was perceived at least by some staff to be playing more of a supporting or cheerleading than a guiding role. This lack of direction was frustrating to some staff members, who would have preferred to skip the process and be told what to do, rather than “going around in circles.” Ongoing provider presence was spottier at PLN and RU schools, as technical assistance for both reforms is an add-on option (or is bundled with AC, but not RU). Both PLN2 and PLN3 had some supervision from PLN staff; at PLN3, this was perceived as helpful, while at PLN2 it was not. At both RU2 and RU3, it was the internal coach (and to some extent at RU2, the district) support to teachers that seemed to play the most important role in fostering or sustaining use, rather than efforts by the external provider per se. At RU3, it appeared that the provider was more of a presence in the school and played a more active role in the early years of implementation.

5.3. Role of Leadership

We have not adopted a formal, theoretically-based definition of leadership for the purposes of this paper, and note that designs differ in regard to whether leadership roles are specified or assumed, as well as to how they are defined. To judge from our data on use of the reforms in schools, however, leadership *in support of the program* divided roughly into five discrete tasks: (1) ensuring that the basic structural conditions existed for program implementation (e.g., time in the schedule, adequate computer hardware); (2) using formal authority to create pressure for change; (3) disseminating knowledge and information about the reform; (4) assisting teachers in making the structural and/or instructional changes called for by the program; and (5) example setting for the program. Within our sample of schools, these five tasks were undertaken by a range of individuals or entities, often with the same person or entity involved in more than one task. In all sites, the presence or absence of a person or persons engaged in these tasks was an explanatory factor for use: Where all five tasks were being addressed and addressed well, use generally exceeded expectations; where none was discernible, use was limited or non-existent.

The task of ensuring basic structural conditions for implementation (Task 1) was undertaken (or not undertaken) by the principal or district. For example, at FTF1 the principal played a central role in creating and supporting the SLCs that are at the heart of the design. Prior to adoption of FTF, he informed the faculty that SLCs would be created, and made clear that anyone who did not want to be part of the process was free to transfer to another school. Once FTF was adopted, he linked the creation of SLCs to his more general goal of improving discipline and school climate.

The use of formal authority to create pressure for change (Task 2) also rested largely with the principal or district personnel. It is worth noting that this task was accomplished (or not) within a broad range of leadership styles. At HSTW2, the principal bluntly stated that teachers not on board with his agenda were not welcome in the school, and deliberately channeled resources and authority to those teachers who he perceived as supporting his vision. At RU2, the principal took a different approach, shining a spotlight on RU teachers as examples to be followed and presenting participation in RU as a desirable activity. Either way, staff at these two schools were given powerful incentives to participate in reform, resulting in relatively higher levels of use at both sites. As discussed in Section 5.6 below, the role of formal leaders, and particularly the principal, helped determine the existence or not of a sense of urgency around a reform.

A wide range of actors were involved in disseminating knowledge and information about the reform (Task 3), including principals, district personnel, provider staff, coaches, School Improvement Facilitators (SIFs), SLC coordinators, focus team leaders, and teachers. For example, at all three FTF sites, SLC coordinators viewed their primary role as disseminating information about the design to teachers in their SLC. The teachers likewise recognized this as their role. Not surprisingly, responsibility for disseminating information about the reform was determined by the reform model. In FTF, it is the task of the SIF and SLC coordinators to pass information to colleagues, while in SN it is the role of the district to disseminate information to schools and teachers. In HSTW the site coordinator and focus team leaders are responsible, while PLN relies on teachers who have taken their courses to spread the word to colleagues through informal communication. With the exception of SN, it appeared that information dissemination tasks were executed at nearly all of our sites.

Several different actors were involved in assisting teachers to make changes called for by the program (Task 4). At SN2, department chairs used SN's web applications to produce classroom-level reports on student performance, which they then provided to teachers for instructional planning. At FTF2, the SIF led professional development on implementing SLCs and facilitating the instructional improvement protocols called for by the design, conducted informal classroom observations to offer feedback to teachers, and sat in on regular SLC meetings. This task is distinguished from Task 3 above in that it often includes but goes beyond the mere dissemination of information. It is more detailed, more individualized, and more directive.

The task of example setting for the program (Task 5), though most frequently embodied in the actions of teachers, was occasionally carried out by coaches and principals as well. At PLN3, for instance, seasoned PLN teachers took on the role of instructors for introductory PLN workshops for colleagues. At HSTW2, the principal created an advisory group comprised of teachers he wished to highlight as positive role models for the rest of the school. Teachers reported that at

least some of their colleagues serving on the committee were influential because they had earned the respect of their peers.

5.4. Zones of Enactment

Another factor that may explain some of the observed variation in use across the five designs is teachers' "zones of enactment":

Zones of enactment refer to that space where reform initiatives are encountered by the world of practitioners and 'practice,' delineating that zone in which teachers notice, construe, construct and operationalize the instructional ideas advocated by reformers (Spillane, 1999, p. 144).

Studying teachers' ways of understanding a mathematics reform initiative and its influence on their practice, Spillane found that even in cases where teachers were aware of or supported reform, the way in which they adopted it seldom challenged or fundamentally altered the "core" of their instructional practice. Specifically, he found that most teachers enacted some of the mechanics of reforms without fully grasping their core ideas, ultimately resulting in superficial change in instructional practice. To explain this discrepancy, Spillane suggested that each teacher has his/her own zone of enactment, shaped by his/her "capacity, will, and prior practice." Teachers whose enactment zones were more "social"—characterized by frequent interaction and deliberation with colleagues and local experts around the reform and availability of resources supporting learning about reform ideas—tended to make deeper, more substantive changes in practice. Those whose enactment zones were more "individualistic"—relying on their own beliefs, knowledge, and experience and less on interaction with colleagues or resources—tended to implement reforms more superficially (Spillane, 1999).

While our data and methodology are not suited for identifying and describing variation across teachers' zones of enactment, the level and type of adoption of external reform among teachers in our schools was very similar to those in the study described above, especially for the designs focused on instruction. Specifically, we found that teachers were often willing and able to adopt procedural aspects of reform, but seldom appeared to grasp the deeper concepts embedded in those procedures. For example, the most prevalent PLN instructional strategies (e.g., pair/small group work, "do now") did not require large departures from typical classroom practice; most were consistent with techniques already used by teachers. On the other hand, PLN strategies with lower rates of use (e.g., using primary source documents, note-taking technique for lectures) asked teachers to more significantly alter their teaching practice. It appears that the further a strategy required teachers to depart from their typical classroom practice, the less likely they were to use the strategy with regularity. Similarly, survey data from RU sites suggest that teachers more frequently used strategies that required little change to their instruction (e.g., guided reading, independent reading) than strategies that further diverged from their current practice, such as teaching specific reading techniques (e.g., decoding skills, word analysis) or conducting reading conferences.

The role of zones of enactment is more complicated with respect to HSTW and FTF. Spillane's (1999) theory suggests that in cases where teachers within a group (e.g., SLC, focus team) work

closely with each other and/or have access to provider resources, one would expect to find deeper, more significant changes in practice. While our data do not allow us to test this hypothesis across all schools or providers, we found such a pattern in at least one HSTW site, and hope to further investigate enactment zones as an explanatory variable in future research. Across the three designs where the reform is intended to reach the entire faculty (FTF, HSTW, and SN), survey data suggest that teachers understand the purpose of the reform, but often do not believe it requires them to make major changes in instruction. A large majority of teachers across all three providers felt they understood the purpose of the design, but 59%, 23%, and 10% of FTF, SN, and HSTW teachers, respectively, indicated that they believed the reforms required them to make major changes in instruction. Given that successful implementation of any of the five designs would result in significant changes in instruction, teachers' ambiguity about this mandate is noteworthy.

5.5. School Resources

The level of and access to three types of school resources—money, time, and materials—also contributed to the observed variation in use both across schools within a design and across the five designs. As described earlier, the five designs call for different levels and mixes of these resources. For example, PLN primarily requires dollars to pay for a small number of teachers to take PLN courses outside of the school day. SN requires teachers to have access to computers and the internet. In contrast, RU and FTF require additional staff to reduce class size (RU) or support school-wide implementation (FTF). Some of the designs, such as RU, FTF, and HSTW, call for schools to reconfigure schedules to make shared planning time and/or more instructional time available.

There were some areas where schools could, and did, implement the designs by reallocating existing resources. For example, districts or schools made professional development funds available to teachers participating in PLN at all three schools, and FTF schools reorganized their staff into SLCs and rescheduled school time to provide for common planning time twice a week. There were areas, however, where the designs required additional resources, including a SIF at FTF schools, additional teachers to reduce class sizes in RU classes, and extensive professional development and/or training of teacher cohorts at PLN3 and at HSTW, FTF, and RU schools. Problems arose when schools could not meet these needs by reorganizing existing resources alone but required additional resources, which districts may not have been willing to or could not provide.

At RU3, larger class sizes and scheduling problems resulted from the school not being able to hire more staff to meet the structure and staffing expectations of RU. Human resource constraints due to finances also affected FTF schools. As a result of budget cuts, FTF2 and FTF3 could not hire enough teachers to maintain pure SLCs, in which students would never to take a class outside of their assigned SLC, as dictated by FTF. In this way, budget cuts affecting human resources (i.e., staffing) in turn affected the structure of the design. Continued funding is also needed to maintain a program's involvement at a school, but consistent funding levels are hard to sustain particularly when schools have depended on external funds, such as federal comprehensive school reform or state school improvement grants, to support the designs. When these grants expired, schools had to look for other funding opportunities which may or may not

align with the design. For example, RU3 was awarded a foundation grant and it was unclear how this different funding stream, with its own goals, would impact the use of RU at the school.

The designs also call for different reconfigurations of time. Both FTF and HSTW require increased amounts of teachers' time to be spent "doing" the many school-wide components of the reforms. The FTF structure requires common planning time during the school day, and this provided teachers with protected time to discuss students, student work, and instruction. Teachers were pressed, however, to find sufficient in- and out-of-school time to implement the FAS component and to develop and prepare for the multiple courses each now offered within their SLCs. Other designs like PLN and SN require additional time to carry out components of the design but do not explicitly call for changes to the structure of the school day. As an instructional reform, PLN requires additional preparation time as well as in-class time to incorporate and use PLN strategies. Class time devoted to using PLN strategies often meant a trade-off between some amount of content and using PLN strategies. This tension between curriculum pacing and content depth was repeatedly noted at both PLN1 and PLN2, and teachers felt uncomfortable with sacrifices made in either direction. At SN3, teachers reported that using Align (i.e., SN's online tool linking standards and curriculum) was extremely time consuming because it required them to enter their lesson plans manually into the database. This frustration was likely compounded by a lack of awareness of the purpose of the tool—teachers viewed it as a repository for lesson plans rather than a tool to align standards and curriculum—and possibly by a lack of proper training, resulting in inefficient use. Either way, teachers reported that Align required substantial investment of time and offered little if any benefit. Across the board, teachers engaged in all designs found themselves needing to do more within the same school day.

Lastly, availability of and access to material resources affected teacher use across the designs. While often linked to financial resources, material resources were an important enough factor in our data to stand alone, especially in the case of SN. Because SN's design is entirely dependent on technology, availability of and access to computers and the internet is requisite. Across all three schools, however, problems related to computer technology (SN2, SN3) and/or the SN program(s) limited and, in some cases, severely hindered use. Examples of the importance of material resources in the other four designs were not as significant as SN's. Nevertheless, they did play a role in how often and in what ways teachers used design components. For example, each RU classroom is required to have an in-class library with leveled books; however, a few RU classrooms did not have a library with the necessary range of books and, as a result, RU teachers could not fully carry out the lessons as prescribed.

5.6. Sense of Urgency

A large body of literature on change, both in schools and in the business world, emphasizes the importance of individuals within an organization believing that change is necessary as a pre-requisite to undertaking it (e.g., Heifetz, 1994; Kotter, 1995). Although we did not specifically ask about the perceived necessity of change in the first rounds of interviews, we did find differences among schools with regard to a sense of urgency around change. The sense of urgency, in turn, appeared to have an effect on use, particularly with regard to early implementing sites. Sometimes, the sense of urgency paralleled the school's compliance status

with state mandates. For example, both FTF1 and RU2 had failed AYP the year of our site visits, there was a detectable sense of the need for change at both sites, and both sites met or exceeded expectations for use for their respective reforms and stages. In the case of PLN2, which had a very high achieving student population, there was almost no sense of the need for instructional change within the building, which was an obstacle for those teachers who were trying to employ PLN techniques in their classrooms. As also found in previous CPRE research, accountability status was not necessarily always an accurate predictor of the existence of a sense of urgency (Goertz & Massell, 2005; Gross & Goertz, 2005). HSTW2, for example, made AYP, yet the principal convinced at least part of the faculty of the need for urgent change, while others were resistant. PLN1 did not make AYP, yet because the school was waiting on the appointment of a new principal, there was little sense of urgency for changes in instruction, and considerable resistance to proposed changes involving a gradual reduction of tracking at the school. Leadership appeared to be a strong intervening variable in determining how a sense of urgency was communicated throughout the school and whether it translated into a positive force for use. This is a theme we wish to look at more closely in our final round of data collection.

6. Conclusion

The interim findings on the use and effects of five externally provided reform models in 15 high schools presented in this paper confirm some lessons learned from studies of implementation of external reforms in the elementary context. In particular, even among schools hand-picked by providers as examples of program use at different stages, there can be variation in use within provider models that exceeds differences attributable to implementation stage (Berends, et al., 2002a, 2002b). Across providers, use varied from almost non-existent to schools which seemed to have wholeheartedly embraced the reform and achieved some measure of both spread and depth. Other findings in line with the existing implementation literature include the greater difficulty of achieving depth over spread of reforms, particularly with regard to achieving changes in instruction, and the very tenuous nature of sustainability. Factors that appeared to explain our findings on use included design emphasis, training strategies, the role of leadership, the concept of teacher zones of enactment, school resources, and the existence or absence of a sense of urgency about reform.

There were a number of areas where our interim data were insufficient to substantiate a hypothesis or pointed in interesting directions but were too thin to present as findings. One such area is the hypothesized existence of a feedback loop between spread and depth, possibly mediated by perceived effects. While we have some evidence, more information is needed about teacher attitudes and beliefs about change and instruction, and what motivates teachers to undertake or sustain a particular reform. We plan to gather additional data on this and other areas when we conduct final site visits in spring 2006. In accordance with the overall study design, we will not return to the mature sites, as these were intended to give a point of reference while the study's main focus is on the dynamics of implementation that are more evident earlier in the process. Finally, the next round of data collection and analysis for the study as a whole will focus on making connections between findings on use and effects from the school perspective with perspectives from the providers and the school districts, as well as with findings on the role of leadership and social networks.

References

- Berends, M., Bodilly, S., & Kirby, S. N. (2002a). *Facing the challenge of whole school reform: New American Schools after a decade*. Santa Monica, CA: RAND Corporation.
- Berends, M., Bodilly, S., & Kirby, S. N. (2002b). Looking back over a decade of whole-school reform: The experience of New American Schools. *Phi Delta Kappan*, 84(2), 168-175.
- Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, 32(6), 3-12.
- Corcoran, T. (2003). *The Merck Institute for Science Education: A successful intermediary for education reform* (CPRE Research Report No. RR-052). Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.
- Datnow, A., Borman, G., & Stringfield, S. (2000). School reform through a highly specified curriculum: Implementation and Effects of the Core Knowledge Sequence. *The Elementary School Journal*, 101(2), 167-191.
- Goertz, M. E., & Massell, D. (2005). *Holding high hopes: How high schools respond to state accountability policies* (CPRE Policy Brief No. RB-42). Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.
- Gross, B., & Goertz, M. E. (Eds.). (2005). *Holding high hopes: How high schools respond to state accountability policies* (CPRE Research Report No. RR-056). Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education.
- Heifetz, R. A. (1994). *Leadership without easy answers*. Cambridge, MA: Belknap Press of Harvard University Press.
- Kotter, J. P. (1995). Leading change: Why transformation efforts fail. *Harvard Business Review*, 72(2), 59.
- Larson, R.L. (1999). *Changing schools from the inside out* (2nd ed.). Lancaster, PA: Technomic Publishing Company.
- McLaughlin, M. W., & Mitra, D. (2001). Theory-based change and change-based theory: Going deeper, going broader. *Journal of Educational Change*, 2(4), 301-323.
- Pascale, R. T., & Sternin, J. (2005). Your company's secret change agents. *Harvard Business Review*, 83(5), 72-81.
- Riggan, M., Supovitz, J. A., & Hovde, K. (2006). *They come in all shapes and sizes: Leaders and high school reform efforts*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

- Shiffman, C. D., Massell, D., Goldwasser, M., & Anderson, J. (2006). *Design as Intended, Design as Enacted: External Assistance Providers and High School Reform*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.
- Spillane, J. P. (1999). External reform initiatives and teachers' efforts to reconstruct their practice: The mediating role of teachers' zones of enactment. *Journal of Curriculum Studies*, 31(2), 143-175.
- Stein, M. K. (2002). *Scaling up mathematics: The interface of curricula with human and social capital*. Proposal to the Interagency Education Research Initiative (IERI). Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.
- Weinbaum, E. H., Supovitz, J. A., Gross, B., Cole, R. P., Weiss, M. J., & Ricalde, B. (2006). *Going with the flow: Communication and reform in high schools*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.