

SEDL's Working Systemically Model

Final Report



RESEARCH REPORT

SOUTHWEST EDUCATIONAL
SEDL
DEVELOPMENT LABORATORY

211 E. 7th St., Suite 200
Austin, TX 78701-3253
800-476-6861
<http://www.sedl.org>

SEDL's *Working Systemically* Model

Final Report
January 2006

Karen S. Herbert, MA
Kathleen M. Murphy, PhD
Miguel A. Ramos, PhD
Michael Vaden-Kiernan, PhD
Joan L. Buttram, PhD



Southwest Educational Development Laboratory
211 E. 7th St., Suite 200
Austin, TX 78701-3253
800-476-6861
<http://www.sedl.org>

©Copyright 2006 by the Southwest Educational Development Laboratory. This publication was produced in whole or in part with funds from the Institute of Education Sciences, U.S. Department of Education, under contract number ED-1-CO-0009. The content herein does not necessarily reflect the views of the U.S. Department of Education, any other agency of the U.S. government, or any other source.

CONTENTS

Acknowledgements.....	i
List of Tables.....	ii
List of Figures.....	v
Executive Summary.....	vi
I. Introduction.....	1
Literature Review.....	3
The <i>Working Systemically</i> Model.....	7
II. Research Design and Methods.....	15
Study Design.....	15
Methods.....	19
Sites.....	19
Data Sources and Methods.....	22
Strengths and Limitations of Data.....	36
Analyses.....	38
III. Field Staff Strategies.....	49
Findings.....	49
The Core Strategies.....	49
Finalizing the Action Plans.....	53
Building Educators' Capacity While Enacting the Action Plan.....	66
Monitoring.....	79
Discussion.....	89
IV. Working Systemically Outcomes.....	91
Findings.....	91
Increases in Systemic Work Over Time.....	91
Differences in Development of Systemic Work.....	100
Exploratory Analyses of Selected Sites.....	113

Discussion.....	123
V. Student Achievement Outcomes.....	126
Supplementary Analyses	127
Findings	128
Student Achievement Results.....	128
Student Achievement and Working Systemically	140
Discussion.....	143
VI. Conclusion.....	145
Findings	145
Recommendations for Systemic Reform.....	147
Recommendations for Research.....	150
References	153
Appendix A: Site Profiles	157
Appendix B: Data Collection Instruments	164
Appendix C: Descriptions of State Assessment Systems	178
Appendix D: Interview Rating Rubrics	189
Appendix E: Complete Data Sources	195
Appendix F: Chi-Square Analyses	200

ACKNOWLEDGEMENTS

The authors would like to thank H. Dickson Corbett and Archie George for their careful reviews of this report as it was developed. We also appreciate the comments from our internal reviewers D'Ette Cowan and Sebastian Wren. All of the SEDL field staff supported our data collection efforts and provided helpful insights. Frances Deviney and Stephanie Bond Huie contributed greatly to our analyses and research design. We are thankful to Debbie Ritenour, Amanda Sealy, and Laura Shankland for their time, patience, and good humor in editing this report. Former SEDL researchers also made valuable contributions in collecting and analyzing data for this project. Finally, we are grateful to our respondents in the districts and schools we visited for their time and candor.

LIST OF TABLES

Table II.1 Project Data by Research Question and Type 19

Table II.2 Participation of SEDL's Intensive Sites 20

Table II.3 SEDL's Intensive Sites 2003–2004 to 2004–2005 21

Table II.4 SCR Entries Across the Stages of the *Working Systemically* Model 26

Table II.5 Fall 2003–Spring 2005 SCR Data Analyzed by Site 27

Table II.6 Fall Interview Sample by Site 28

Table II.7 Sample of Fall 2003–2004 Interview Respondents 29

Table II.8 Spring 2005 Interview Sample by Site 29

Table II.9 Sample of Spring 2005 Interview Respondents 30

Table II.10 *Working Systemically* Survey Sample by Site, 2003–2005 31

Table II.11 *Working Systemically* Survey Sample Respondents, 2004–2005 32

Table II.12 Overview of Student Achievement Data 34

Table II.13 Schools Included in Student Achievement Analysis 35

Table II.14 Sample Chi-Square Table for Third-Grade Reading in SEDL School A 46

Table III.1 Most Common Areas of Focus for Site Action Plans, 2004–2005 55

Table III.2 Time Spent Prior to Stage 4, Taking Action and Monitoring Implementation 56

Table III.3 Other School-Level Initiatives Across the Sites 63

Table III.4 Site-Based Teams That Promoted Leadership Capacity 67

Table III.5 Tools for Monitoring Implementation of Action Plans 86

Table IV.1 Descriptive Statistics for Interview Scores, 2003–2004 92

Table IV.2 Descriptive Statistics for Expectations and Priorities Rubric 93

Table IV.3 Descriptive Statistics for Alignment Rubric 94

Table IV.4 Descriptive Statistics for Collaboration Rubric 95

Table IV.5 Descriptive Statistics for Professional Development Rubric 96

Table IV.6 Descriptive Statistics for *Working Systemically* Survey, 2003–2005 98

Table IV.7 Univariate ANOVA on Interview Scores by Site and Year 101

Table IV.8 Univariate ANOVA on Alignment Scores by Site and Year..... 102

Table IV.9 Univariate ANOVA on Collaboration Scores by Site and Year 102

Table IV.10 Descriptive Statistics for Test/Development Interview Scores..... 104

Table IV.11 Univariate ANOVA on Interview Scores by Test/Development and Year 105

Table IV.12 Descriptive Statistics for Test/Development *Working Systemically* Survey Scores 105

Table IV.13 Univariate ANOVA on *Working Systemically* Survey Scores by
Test/Development and Year 106

Table IV.14 Descriptive Statistics for Interview Scores by Role 107

Table IV.15 Univariate ANOVA on Interview Scores by Role and Year 108

Table IV.16 Descriptive Statistics for Expectations and Priorities Scores by Role 108

Table IV.17 Descriptive Statistics for Alignment Scores by Role 109

Table IV.18 Descriptive Statistics for Collaboration Scores by Role..... 109

Table IV.19 Descriptive Statistics for Professional Development Scores by Role 109

Table IV.20 Univariate ANOVA on Expectations and Priorities Scores by Role and Year..... 110

Table IV.21 Univariate ANOVA on Professional Development Scores by Role and Year 110

Table IV.22 Rankings of Sites by Final Year Mean and Change Scores..... 114

Table IV.23 Significant Changes on *Working Systemically* Survey Items–Wrightsville, OK.... 121

Table V.1 Summary Table of Achievement Across Sites..... 129

Table V.2 Summary Table of Significant Changes Over Time..... 130

Table V.3 Summary Table of Significant Differences Between SEDL School Results and
Comparison Group Results..... 130

Table V.4 Chi-Square Analysis, Grade 6 Math, AR, Grisham–Middle School A 133

Table V.5 Chi-Square Analysis, Grade 5 Reading, OK, Wrightsville–Elementary School B 134

Table V.6 Chi-Square Analysis, Grade 3 Reading, TX, Pineland–Elementary School A..... 135

Table V.7 Chi-Square Analysis, Grade 8 Math, LA, Athens–Middle School A..... 136

Table V.8 Chi-Square Analysis, Grade 8 Math, LA, Bayou City–Middle School B 138

Table V.9 Chi-Square Analysis, Grade 4 Reading, TX, Pineland–Elementary School A..... 139

Table V.10 Correlation Between Interviews and Achievement, 2003..... 141

Table V.11 Correlation Between Working Systemically Measures and Achievement, 2004..... 141

Table V.12 Correlations Between Achievement and Interview Subscales, 2003 142

Table V.13 Correlations Between Achievement and Interview Subscales, 2004 142

LIST OF FIGURES

Figure I.1 SEDL’s *Working Systemically* Model..... 9

Figure II.1 SCR Data Sampling Strategy 24

Figure III.1 Field Staff Strategies During Taking Action and Monitoring Implementation 51

Figure III.2 Action Plan Template 54

Figure III.3 Professional Teaching and Learning Cycle (PTLC) 59

Figure III.4 Sample Action Plan Implementation Review 84

Figure III.5 Sample Innovation Configuration 85

Figure IV.1 Increase in Interview Scores, 2003–2004..... 99

Figure IV.2 Increase in *Working Systemically* Survey Scores, 2004–2005 99

Figure IV.3 Change in Expectations and Priorities Scale by Role 112

Figure V.1 Grade 6 Math, AR, Grisham–Middle School A..... 132

Figure V.2 Grade 5 Reading, OK, Wrightsville–Elementary School B 133

Figure V.3 Grade 3 Reading, TX, Pineland–Elementary School A 134

Figure V.4 Grade 8 Math, LA, Athens–Middle School A 136

Figure V.5 Grade 8 Math, LA, Bayou City–Middle School B..... 137

Figure V.6 Grade 4 Reading, TX, Pineland–Elementary School A 138

EXECUTIVE SUMMARY

Since December 2000, the Southwest Educational Development Laboratory (SEDL) has developed and refined a systemic model to improve student achievement in reading or mathematics in sites composed of low-performing districts and schools (Southwest Educational Development Laboratory, 2000). The Charles A. Dana Center at the University of Texas at Austin partnered with SEDL on this work, with support from American Indian Research and Development.¹ Funding was provided through a 5-year research and development (R&D) contract awarded by the U.S. Department of Education.

This report describes the *Working Systemically* model developed under this contract, presenting analyses of data related to its impact in 12 sites distributed across SEDL's five-state region (Arkansas, Louisiana, New Mexico, Oklahoma, and Texas). It focuses on the final 2 years of the project (2003–2004 and 2004–2005), documenting the sites' progress in developing systemic work and improving student achievement.

Research Questions

The goal of the final report is to describe the progress of SEDL's sites in working systemically as conceived by the model and to explore the degree to which this work is related to student achievement. Specifically, the report addresses the following three research questions:

- What strategies did SEDL field staff use to build the capacity of low-performing districts and schools to work systemically?
- To what extent did low-performing districts and schools increase their capacity to work systemically?
- Did student achievement increase as districts and schools increased their capacity to work systemically?

¹ Throughout the remainder of this report, "SEDL" is used to designate the partnership of organizations that are actively involved in this work.

Data and Analysis

To gather data on the intensive site work undertaken by field staff and on the progress of the sites to build capacity to work systemically and improve student achievement, the research team

- interviewed a sample of educators at each site in the fall and spring;
- administered a paper-and-pencil survey at each site;
- examined site contact records maintained by field staff; and
- analyzed state-mandated achievement test results.

For data describing the field staff strategies, research staff analyzed field staff's site contact records and reviewed exit interviews conducted in Spring 2005 to triangulate these analyses. For outcome data regarding capacity for systemic work, the study included surveys of site educators from 2003, 2004, and 2005 (conducted in the spring semesters), and fall interviews from 2003 and 2004. Student achievement data for 2003 and 2004 were collected for schools across sites. When possible, data were also obtained for 2005. In addition, demographic and achievement variables were used to develop matched comparison groups for each grade-level analysis of achievement.

Findings

The question of whether systemic work is linked to increased student achievement highlights the multiple demands intrinsic to systemic reform, particularly in low-performing districts. Fundamentally, districts must know how to recognize critical needs (e.g., improving achievement) and be able to devise and implement plans to meet those needs. A tacit assumption of this process, however, is that all stakeholders have the capacity to be fully engaged. The wide variation in our results suggests that low-performing districts require more assistance to support their efforts toward building coherence.

Field Staff Strategies Focused on Instructional Leadership

Field staff strategies to develop educators' capacity for systemic work were examined, with particular focus on those designed to have the most direct impact on the classroom. Three core strategies were central. One was to work with site staff to finalize written action plans and ensure that all initiatives in the site were coherent with its goals. A second was to build educators' capacities. To develop site leaders' capacity, field staff relied primarily on collaborative work within structured leadership team meetings. They also offered leaders informal coaching. To strengthen classroom-level capacity, field staff engaged educators in the Professional Teaching and Learning Cycle (PTLC), a collaborative approach to classroom practice that catalyzes the alignment of instruction with curriculum, assessment, and state standards. A third core strategy was to help leaders in the sites to monitor two areas: activities related to the PTLC and progress on related goals at the district level. Despite the emphasis these three core strategies placed on classroom-level practice, all of them stressed the importance of instructional leadership. The accountability of educators to the levels above their own proved crucial to conducting the work supported by the field staff.

Limited Growth in Systemic Work

Sites made progress in working systemically. They developed skills and practices that enable them to integrate and direct the various facets of their school systems toward achieving student learning goals. Important practices in which improvements were made included focusing district priorities and expectations of teachers on improving student achievement, aligning curriculum with state standards, supporting collaboration among teachers on curriculum and instruction, and engaging in high-quality professional development focused on student learning needs. However, the amount of growth found was generally small, and there was variation in outcomes among sites and among groups of respondents. Teachers and administrators diverged on many measures, including the extent to which student learning was made a priority and the quality of professional development in their districts. Differences in patterns of participation in

SEDL's work and knowledge of new initiatives resulting from the work most likely contributed to these differences. Whatever the reason, these disparities among educators in the sites made it less likely that they would be able to exhibit more dramatic growth in systemic work.

Mixed Student Achievement Patterns

The limited gains in working systemically may have affected student outcomes as well. Overall, grade-level achievement results did not indicate a pattern of significant change among schools participating in the project. Thus, the initiative did not provide compelling evidence of impacts on long-term criterion-referenced indicators of academic achievement. However, significant correlations between working systemically measures and academic achievement, particularly for alignment, provide non-definitive but encouraging evidence that progress on these short-term outcomes may be important in regard to academic achievement.

Recommendations for Systemic Reform

The process of systemic reform is two-fold. First, reform efforts have to focus on increasing school and district capacity, including the ability to recognize critical needs (e.g., improving achievement) and to devise and implement plans to meet those needs. Second, program personnel have to help educators utilize capacity coherently and effectively at various system levels. In particular, improvement initiatives need to focus explicitly on skills related to classroom instruction and developing school and district infrastructures to support and sustain coherence in instructional programs.

Increasing School and District Capacity

Both field staff and educators in the sites reported that they believed more time to develop systemic work would be necessary in order to affect student achievement. Our analyses suggest that one reason may be that the implementation of improvement initiatives requires knowledge, skills, and resources that schools and districts do not always have.

- A variety of resources are necessary to support systemic reform. Among these is ensuring sufficient time for establishing processes that contribute to strengthening an educational system. Such processes include collaborative planning and decision making, as well as ongoing professional development.
- Many sites lacked instructional materials that were aligned with their states' standards. Textbooks, planning guides, and other curricular materials should be aligned to and focused on state standards in order to be useful in creating coherence in a district's instructional program.
- Some school principals also found it challenging to incorporate practices related to instructional leadership alongside their operational and managerial responsibilities.

Helping Educators Utilize Capacity

Educators benefit from ongoing coaching as they implement systemic reform efforts. Growth in capacity should be focused on supporting initiatives to align curriculum, instruction, and assessment to state standards, with particular emphasis on instruction.

- In putting the *Working Systemically* model into practice, field staff focused on systemic processes such as collecting data, setting goals and objectives, and planning initiatives. Many educators, however, had had little experience in working collaboratively and needed training on how to use their planning time productively to achieve explicit goals. Field staff therefore emphasized coaching and modeling in their work to build educators' capacity while implementing new programs and initiatives being developed through districts' work with the *Working Systemically* model.
- Alignment was positively and consistently related to improved achievement. Directly addressing curriculum and instruction may lead to improved student learning since

alignment was the area of practice most directly related conceptually to teaching and learning.

- As the *Working Systemically* model suggests, improvements to curriculum and instruction do not occur in isolation from other school and district practices. Other areas of the system must support such improvements and should be directed toward improved teaching and learning. Staffing practices, for example, must focus on recruiting and retaining teachers who have the content and pedagogical knowledge to implement instructional strategies that meet students' needs. Administrators also need to be able to evaluate the quality of instruction and to support teachers in improvement efforts.

Recommendations for Research

The recommendations for research that follow relate both to content and methodological concerns. Generally, they suggest that researchers have yet to address adequately the capacity issues that emerged in this work. Following a list of three suggested topics for future research are some methodological recommendations regarding research in low-performing educational systems.

Areas for Further Research

Future research on systemic reform should focus on designing dual-focused technical assistance, determining specific impacts of alignment on student achievement, and doing more systematic classroom observations.

- Because systemic reform is by its nature multilevel, it calls for a dual focus in the kinds of technical assistance provided. Some teachers and school-level administrators, however, reported feeling peripheral to SEDL's work until it more directly engaged issues of instruction. Researchers need to examine more

systematically what types of technical assistance will support administrators' leadership capacity as it simultaneously improves instruction.

- Curriculum and state standards, as well as state-mandated assessments, are all explicitly connected. The “black box” in the alignment process is instruction. Future research should elaborate upon exactly why alignment might boost student achievement (Newmann et al., 2001).
- Future research needs to emphasize more formal observations of classroom practice that will systemically document which instructional strategies convey a standards-based curriculum in ways that will improve students' performance on aligned assessments. Researchers can also focus on how best to ensure that teachers know how to implement these strategies.

Methodological Considerations

The majority of sites participating in this work were selected due to low performance on state standards-based achievement tests. However, this focus introduced other issues particularly endemic to these environments: limited knowledge of what engagement in a research project entails, concurrent school improvement programs, and administrator turnover.

- Knowledge of what participation in a research project entails is generally limited to scholars. Any school improvement intervention to be evaluated scientifically must therefore incorporate strong incentives for participants, leaders especially, to ensure continued commitment to it. This process of recruitment has to be ongoing.
- Particular attention needs to be paid to the external environment in which work in low-performing districts and schools is conducted. These systems often utilize a number of school improvement programs in efforts to improve student achievement within the context of high-stakes accountability. Educators and field staff in this

project often had to negotiate among the demands of several of these programs. Such negotiations can lead to variations in the ways programs are implemented, which must be documented and accounted for in evaluations and research.

- Changes in leadership can greatly influence the direction of a reform process positively or negatively. In either case, a period of adjustment necessarily follows. Ideally, fostering an awareness among leaders of the role they play in the success of research interventions would be part of the coaching mentioned above.

Implications

The question of whether systemic work is linked to increased student achievement highlights the multiple demands intrinsic to systemic reform, particularly in low-performing districts. If systemic work is a goal for substantially low-performing districts, the initial phase of the program must focus resources toward helping educators at all levels attain a threshold of competency in needed areas. Long-term success is largely dictated by how well all the parts of the educational system work toward a common goal. Continued improvement is only possible when educators at all levels have the capacity to engage effectively in the reform process.

I. INTRODUCTION

Since December 2000, the Southwest Educational Development Laboratory (SEDL) has developed and refined a systemic model to improve student achievement in reading or mathematics in sites composed of low-performing districts and schools (Southwest Educational Development Laboratory, 2000). The Charles A. Dana Center at the University of Texas at Austin partnered with SEDL on this work, with support from American Indian Research and Development.² Funding was provided through a 5-year research and development (R&D) contract awarded by the U.S. Department of Education.

This report is the fourth in a series that describes the *Working Systemically* model developed under this contract, presenting analyses of data related to its impact. The first report (Southwest Educational Development Laboratory, 2002) outlined the site selection and recruitment procedures. It also provided demographic and other descriptive information about the sites, which were located in districts across SEDL's five-state region (Arkansas, Louisiana, New Mexico, Oklahoma, and Texas). The second report (Southwest Educational Development Laboratory, 2003) summarized research progress through December 2002. By this time, SEDL had begun early work in participating sites. The report identified major factors that had an impact on the model's development. During the 2003–2004 school year, SEDL refined the model. Site work began to emphasize the alignment of four of the model's components: standards, curriculum, instruction, and assessment. The third report (Southwest Educational Development Laboratory, 2004) presented data relevant to understanding how well educators in the sites understood alignment conceptually. It also described field staff activities in the sites and presented student outcome data available for the sites at that the time. During the final year, 2004–2005, site work concentrated on improving classroom instruction while continuing the processes of aligning curriculum and assessments to standards. This final report describes field

² Throughout the remainder of this report, "SEDL" is used to designate the partnership of organizations that are actively involved in this work.

staff strategies for using the *Working Systemically* model to build systemic work in the sites during the final 2 years of the project (2003–2004 and 2004–2005). It also documents the progress of the sites in terms of developing systemic work and improving student achievement.

This introduction (Section I) provides a brief review of the school reform literature, SEDL's *Working Systemically* model, and an overview of the stages of the model. Section II presents the research design for this report. This section includes the research questions; an overview of the research sites; a discussion of the data sources and their strengths and limitations; and a description of the methodological procedures used for each research question's analyses. Sections III, IV, and V present the findings and discussions related to each of the three research questions. Section III analyzes what strategies field staff used to initiate systemic reform of the educational systems in the research sites. Section IV explores sites' growth on measures of working systemically. Section V examines patterns in student achievement data available for tested grades in the participating schools. The report concludes (Section VI) with an analysis of the relationship between capacity for systemic work and measures of student achievement. The reader will find site profiles in Appendix A, which contain in tabular form detailed site demographic information. The instrumentation SEDL used to gather the data for this report is in Appendix B. Appendix C describes the assessment systems used in each of the five states in which SEDL worked. Appendix D contains a set of rubrics used in Section IV to measure systemic work as expressed in interview data. Appendix E includes tables describing the universe of data collected during this 5-year project, some of which were analyzed in previous reports. In Appendix F appear contingency tables and chi-square (χ^2) statistics for grade-level analyses presented in Section V. Statistics are presented after each table.

Literature Review³

The use of standards and assessments to improve student achievement has been an important policy tool in public education for over 2 decades (Puma et al., 2000). However, districts and schools often lack the organizational skills and infrastructure to implement standards-based reforms in a coherent, sustainable manner. SEDL's *Working Systemically* model was designed to assist low-performing districts and schools in such systemic improvement.

Beginning with research on school effectiveness, researchers have identified many features associated with successful schools. Clear goals for student learning, parent involvement, and strong principal leadership have all been found to be common features of high-performing schools (Edmonds, 1979; Hallinger & Murphy, 1986; Purkey & Smith, 1983). Research on private schools has also revealed that shared values, communal climates, and an emphasis on academic subjects contribute to effectiveness in achieving high academic outcomes (Bryk et al., 1993). Additionally, studies of school restructuring have shown that shared decision making, collaboration among teachers, and clear goals for student learning are critical to improved student learning (Bryk et al., 1997; Newmann & Wehlage, 1995).

These lines of research have produced useful descriptions of high-performing schools and have identified the conditions and features that allow them to flourish. However, questions remain about how to transform low-performing schools into high-performing ones and how to implement practices in schools and districts that will enable them to sustain success. In the context of standards-based reform, these questions are particularly salient as schools and districts face increasingly rigorous standards for performance and loftier goals for student achievement. While research has identified strategies useful to policymakers in implementing standards-based

³ While this report attempts to ground SEDL's *Working Systemically* model in the research and theory of the school reform movement, it is not intended to be a comprehensive review of the reform literature. We invite interested readers to utilize the citations provided in this brief review for further information on the history of school reform.

reforms (Massell, 1998; O'Day et al., 1995), less is known about strategies and practices for implementation of these reform models at the school and district levels.

Most school improvement models designed to achieve the level of performance envisioned in standards-based reform emphasize either improvement of the school's curriculum and instructional programs or the organization of the school itself. These models have contributed greatly to educators' understandings about practices and structures that lead to improvement in certain areas of school performance but are insufficient for coherent, sustainable reform.

Reform of Curriculum and Instructional Programs

Numerous school reform efforts have concentrated on implementing curricular and instructional innovations (Datnow et al., 2000; Sizer, 1996; Slavin et al., 1996). A significant contribution of these reforms is the attention placed on content knowledge, pedagogy, and alignment, all of which directly impact student achievement. Specifically, Datnow and colleagues (2000) noted that the alignment of curriculum with instruction leads to greater collaboration between teachers and decreased curricular repetition between grades.

A major limitation of this approach, unfortunately, is that educational systems may have neither the capacity to implement the selected curriculum and instruction program nor the supports necessary to sustain improvements over time (Fullan, 2001). For these reasons, practitioners may respond to top-down reforms, such as state or district policy initiatives, with resistance or confusion (Cohen & Ball, 1990) and policymakers and administrators may have to ensure that adequate resources and curricula are available (O'Day et al., 1995; Sashkin & Egermeier, 1993).

Reform of the School Organization

Another reform approach addresses problems at the school organizational level. Many organizational reform efforts have drawn their theories and strategies from successful business models. An example is school-based management, which assumes that the closer decision makers are to students, the better the decisions will be (Leithwood & Menzies, 1998). These types of reforms are valuable in focusing attention on the organizational structures that enable schools to respond effectively to their particular problems and populations, which will in turn increase the potential of schools to perform at higher levels in the future (Honig & Hatch, 2004). The organizational lens also emphasizes sustainability of improvements over time (see Bryk et al., 1997; Newmann & Wehlage, 1995).

As with curricular and instructional reforms, a limitation of the organizational approach is that it often does not address how districts and states impact decisions made at the school. Fullan (2001) describes the problem of not addressing all levels of the educational system as follows:

The main reason that change fails to occur in the first place on any scale, and does not get sustained when it does, is that the infrastructure is weak, unhelpful, or working at cross purposes. By the infrastructure I mean the next layers above whatever unit we are focusing on. In terms of successive levels, for example, . . . a school can initiate and implement successful change, but cannot sustain it if it is operating in a less than helpful district; a district cannot keep going if it works in a state which is not helping to sustain reform (p. 18).

Additionally, although organizational improvement models allow for each school to interpret reforms as best fits their needs, inequities between schools in resources, educator expertise, or effective instructional methods may still hamper their efforts (Leithwood & Menzies, 1998). Finally, organizational reform efforts often include issues that are tangentially rather than directly related to student achievement.

An Integrated Approach to Reform: Systemic Improvement

As accountability systems have become more comprehensive, educational research has increasingly turned to systemic approaches to improvement that encompass both instructional practice and organizational reforms (see Corcoran & Christman, 2002; Newmann et al., 2001; J. P. Spillane & Thompson, 1997). This approach emphasizes developing the capacity of schools to meet higher standards and expectations for performance. While capacity remains an elusive concept whose definition often depends on the issue or group to which it is being applied (Cohen & Ball, 1999), in the context of school improvement it can be thought of as the knowledge, skills, and supports needed to develop and utilize resources effectively toward improving student outcomes (Hatch, 2001).

These conceptions of capacity include not only the knowledge, skills, and motivations of individuals within a system, but also the collective resources contributed by schools, districts, and the state toward reform goals (Massell, 1998). Recent research on systemic initiatives has pointed to district support, school and district leadership, teachers' knowledge of standards and curriculum, and professional development as areas in which capacity is needed in order to improve instructional and organizational practices (Corcoran & Christman, 2002; Goertz & Massell, 2005). Spillane (2005) also found that schools and districts often misinterpret and misunderstand state policies related to accountability and standards-based reforms, resulting in ineffective implementation. Understanding of standards and accountability policies is another area in which capacity is needed in order to implement systemic reform and improve student achievement.

While most research on this approach remains theoretical in nature, some studies have examined the effects of systemic improvements on student achievement, with mixed results. An evaluation of the *Children Achieving* systemic reform program in Philadelphia, for example, found only a few modest gains as a result of the program (Corcoran & Christman, 2002). The

lack of results were attributed to problems with the program's theory of action, uneven implementation, lack of capacity for the reforms, and lack of support from teachers.

Newmann and colleagues (2001), however, found that elementary schools with coherent instructional programs had higher performance on standardized tests in reading and math. The authors defined a coherent instructional program as one that was "guided by a common framework for curriculum, instruction, assessment, and learning climate and that [was] pursued over a sustained period," (p. 297). This emphasis on coherence and alignment to goals and standards for student learning are central to the systemic approach and are also key elements to SEDL's *Working Systemically* model.

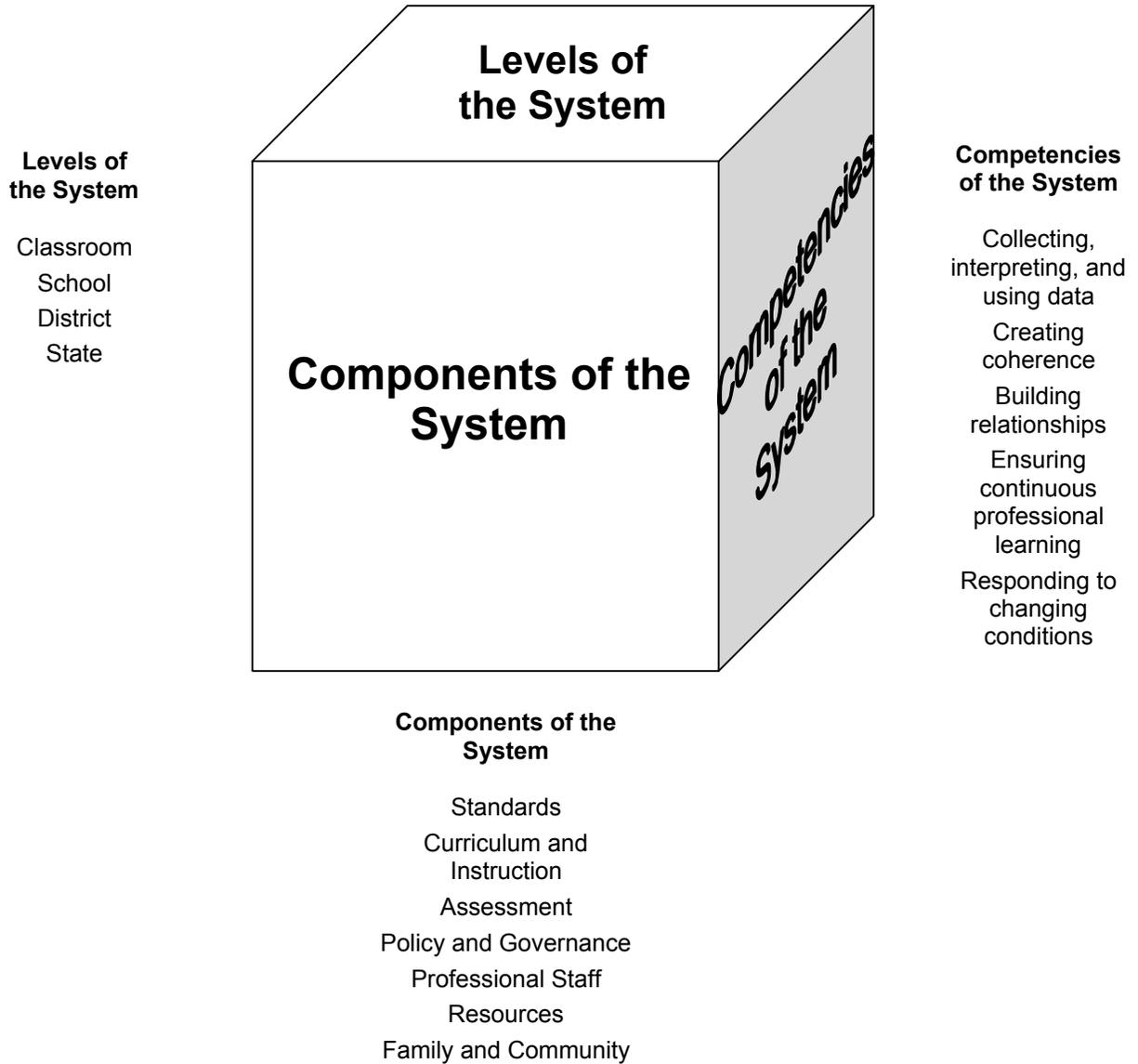
The *Working Systemically* Model

Failure to take into account how the factors that affect the reform of an educational system interact with one another limits the impact of many reform models. Improving the function of any one academic component (e.g., curriculum) or organizational level (e.g., principal's leadership) in isolation will have minimal impact (Fullan, 2001). Systemic approaches address how the academic components and organizational levels work together to initiate and sustain change over time (Smith & O'Day, 1990). Each level of a system, for example, holds the level below it accountable for implementing initiatives, while the level below depends on the one above for necessary resources and direction.

SEDL's *Working Systemically* model draws on the strengths of the curricular and organizational reform models and emphasizes the interactive dimensions of educational systems. The model provides a three-dimensional framework that represents the factors that educators must address as part of their reform work (see Figure I.1). First, working systemically involves all levels of the educational hierarchy (i.e., classroom, school, district, and state). Second, all components that directly impact education must be considered in making improvements (i.e., standards, curriculum and instruction, assessment, policy and governance, resources,

professional staff, and family and community). Third, educators must master five competencies so that they have the needed skills to achieve and sustain reform: (1) collecting, interpreting, and using data; (2) creating coherence among all items on these dimensions; (3) building relationships; (4) ensuring continuous professional learning; and (5) responding to changing conditions. The educational system must effectively address each dimension to understand how to design and implement broad scale improvement initiatives (see Southwest Educational Development Laboratory, 2000, for an in-depth discussion of the *Working Systemically* model).

Figure I.1 SEDL's *Working Systemically* Model



Educators in most schools and districts, much less low-performing ones, typically are not equipped with the knowledge and skills necessary to manage the complexities of systemic change (Center for Mental Health in Schools, 2001). Thus, external facilitators can be helpful catalysts to substantive reform across an educational system. They can bring additional skills to the site, transfer them to educators, and support the establishment of new structures created to maintain improvement efforts over time (Fullan, 2001). SEDL addressed both the technical needs and the organizational structures of each of the participating sites in the project by

providing two facilitators who supported educators as they promoted systemic work as described in the *Working Systemically* model and its associated practices and values.

One of these two facilitators, the site coordinator, was often a former administrator with experience in successful school reform. This person had responsibility for overseeing the process of change in a given site. Site coordinators focused especially on developing leadership skills so that administrators knew how to allocate resources based on priorities related to student achievement. Site coordinators also encouraged leaders to engage in collaborative decision making that promoted shared responsibility and, consequently, coherence of effort. The second facilitator was the site specialist, who was either a reading or math specialist. The site specialist helped educators analyze student achievement data to identify areas of strengths and weakness in the district's instructional program. They then assisted in designing curricular, instructional, and assessment strategies to boost student achievement. The site coordinator and site specialist treated different aspects of the model, but they worked closely together to strengthen systemic competencies to align components throughout all levels of the system. Given this mutual support, in the overview description of the stages of the model that appears in the next section below, reference is made only to the field staff generically, rather than specifying a particular role.

Consistent with the collaborative approach modeled by the site coordinator and site specialist through their partnering, field staff established leadership teams at each site to initiate and carry out the model's implementation. These teams comprised cross-sections of representatives from both district and school levels. Team members directed the site's progress through the model's five stages: (1) Understanding the System, (2) Analyzing the System, (3) Planning Action, (4) Taking Action and Monitoring Implementation, and (5) Assessing and Reflecting on Outcomes.

SEDL conducted this work through the end of the academic year 2005 in five states (Arkansas, Louisiana, New Mexico, Oklahoma, and Texas). Twelve sites were engaged with the

project at its completion in May 2005. SEDL assigned a site coordinator and a reading or math site specialist to each site. These field staff were responsible for helping district- and school-level staff use the *Working Systemically* framework and for developing a set of protocols to structure and direct sites' work. The field staff initiated their interactions with the leadership teams in a "guide and model" mode to shepherd them through the stages. Eventually, each leadership team was to take on more of the responsibility for planning and directing the strategies for systemic work and improving student achievement as a necessary prelude to sustaining the changes made.

Stages of the Working Systemically Model⁴

The first three stages of the *Working Systemically* model take a site through a process of identifying its educational needs via collaborative efforts engaging district and school leaders in collecting and analyzing data about the site and its students. Together with the field staff, the team members decide which needs to address first and then write plans that specify how to improve in these areas. Stage 4 concerns enacting the plans and monitoring associated activities and processes, whereas stage 5 addresses the assessment of outcomes related to the success of the processes related to systemic work and student achievement. During this final stage, participants consider whether their stage 3 objectives have been met and, if so, celebrate their accomplishments. If not, they reflect upon what hindered their progress. In either case, educators review data to consider how best to move forward with renewed effort. Initially, the field staff direct the work. After a site goes through all five stages, educators should take on more leadership.

⁴ This section is adapted from drafts of the *Working Systemically* model's toolkit, developed by field staff based on their experiences. In effect, the toolkit represents the culmination of years of effort to develop and refine not only the model but also its protocols for implementation. In earlier versions of the protocols, the stage now called Understanding the System was represented by two stages: Site Entry and Data Scan. Readers who compare this description of the model's stages to those in prior reports may also notice a slight rewording of stage 3's title from Systems Exploration to Analyzing the System. Another difference is that stage 4, Taking Action and Monitoring Implementation, has been renamed. Its prior title was Taking Action and Monitoring Results. Assessment of results, field staff found, was more appropriately situated within stage 5, which was accordingly renamed Assessing and Reflecting on Outcomes. Previously, it was called Recycling for Continuous Improvement. Otherwise, the descriptions of the stages in this report do not substantially differ from what is found in prior publications on this project's data.

A first step associated with stage 1, Understanding the System, is for the field staff to present the model to the site. Site leaders are an especially important audience at this stage because their commitment to the work is crucial to ensuring accountability to it. At this stage, schools are most involved in data collection efforts, which focus on student achievement, demographics, and current instructional practices in core subjects. Data are also gathered and analyzed regarding how well the site's curriculum, instruction, and assessment are aligned with state standards. The methods leaders in the site use to monitor and support student achievement are closely examined. The teams generate reports that make recommendations on whether the site should move forward with improvement, the school(s) where the work should occur, and the content focus (e.g., reading, math). The teams also share the reports with schools and all parties involved, who by the end of this stage commit formally to the work and sign a Memorandum of Understanding. Without this signed memorandum, the work should not continue.

In stage 2, Analyzing the System, the field staff and administrators select other members of the district leadership teams, typically principals, content specialists, teachers, and other individuals who hold formal or informal leadership roles in the system. These expanded district leadership teams review district requirements for Annual Yearly Performance (AYP) and trends in student achievement and then administer and analyze surveys of schools regarding the status of their instructional programs in the selected content areas. Based on these data, the members create a "problem statement" that highlights key "gaps" between the system's actual situation and the ideal state described in research. These gaps are areas on which the site should focus. The team identifies and prioritizes critical elements that must be addressed to reach the ideal state. The team then presents these descriptions of the current and ideal states and consequent priority areas for improvement to school representatives. The district leadership team asks for schools' ideas about strategies for aligning curriculum, instruction, and assessment with the state standards; working collaboratively and scheduling school time to support improvement; and achieving the district goals at the school level. The team reflects on the comments and

subsequently revises and/or confirms district priorities. By the completion of this stage, leaders have created a set of goals that will serve as the foundation of a site action plan.

Stage 3, *Planning Action*, produces this action plan. The district leadership team investigates research-based actions essential to impacting classroom instruction and improving student achievement. They focus especially on identifying processes that increase classroom-level alignment of curriculum, instruction, and assessment with state standards. The team also discusses what role leaders will play in ensuring sustained commitment to the effort across the levels of the system and work with leaders in participating schools to develop a detailed plan for accomplishing the goals set in the previous stage. Stage 3 concludes with formal adoption of the plan and a firm commitment from leaders at all levels to allocate the requisite time, funds, and personnel as outlined in the plan.

During stage 4, *Taking Action and Monitoring Implementation*, staff members carry out the planned actions as scheduled in the district and school timelines at the district, school, and classroom levels of the educational system. District and school leaders both formally and informally monitor the steps recently taken, evaluate the impact of those steps, and review the planned next steps to determine if modifications need to be made to the action plan and/or timeline. As leaders make the agreed upon changes, they ensure that staff members have a clear understanding of the expected next steps, the resources necessary to carry out the required actions, and the knowledge and skills to carry out the work for which they are responsible. Staff members then take action and the cycle continues.

This shift from a linear to cyclical action is the first of three transitions that occur in this stage. The second is that district and school leaders assume more ownership and active management of the process. The third is that there is an increased emphasis on school and classroom levels as leaders initiate new processes, procedures, and approaches to classroom instruction. This stage of work is where the number of advocates for the work increases and long-term sustainability becomes a formal consideration.

Stage 5, *Assessing and Reflective on Outcomes*, is the last of the five stages in the *Working Systemically* model. In this stage of the work, the field staff guide the members of the district leadership team through a reflective process where they assess the progress that has been made to date and the outcomes of the work. This reflection enables the team members to make more informed decisions about their focus and starting points for the upcoming year. The team outlines examples of evidence to consider at this point. Such evidence might include student achievement trends, data from classroom observations, and the discussions, input, and decisions that took place at meetings during stage 4. The field staff summarizes the team's conclusions in a report, at which time members determine whether the system implemented the plan with fidelity. They then review to what degree the site has achieved the intended goals and target outcomes, identify challenges that staff members encountered along the way to be remedied in the future, and recognize and celebrate accomplishments.

II. RESEARCH DESIGN AND METHODS

Study Design

The goal of the final report is to describe the progress of SEDL sites in working systemically as conceived by the model and to explore the degree to which this work is related to student achievement. Specifically, the report addresses the following three research questions:

- What strategies did SEDL field staff use to build the capacity of low-performing districts and schools to work systemically?
- To what extent did low-performing districts and schools increase their capacity to work systemically?
- Did student achievement increase as districts and schools increased their capacity to work systemically?

As Erickson (2002) pointed out, it is important to describe an intervention “as actually delivered” before asking, “Did it work?” (p. 21). The first question this report raises is consistent with this concern. Before analyzing the *Working Systemically* model’s outcomes, the report first describes field staff strategies to develop systemic work in the sites.⁵

A committee of the National Research Council noted that various kinds of questions call for different kinds of methods within social science:

What is happening (description); is there a systematic effect (cause); and why or how is it happening (process or mechanism)? A range of methods can legitimately be employed to address each type of question, and the choice should be governed by the particular purposes and circumstances of the research (Feuer et al., 2002, p. 7).

The methods for the first question fit into the first query on this list (“description”), with some elements of the third (“process”), given that in the course of this description, the reader

⁵ In the context of this report, the term “site” refers to the group of participating schools within a given district. In a practical sense, the use of “site” vs. “district” was a semantic choice because the number of sites is equivalent to the number of districts involved in the project. However, not all schools in a district were working with SEDL. Although a district-level process was implicit in the *Working Systemically* model, “site” more accurately reflects the degree of direct involvement with schools in each district.

should gain an understanding of how the field staff intervened at different levels across the sites. The methods proposed facilitate the production of “rich descriptions of the nature of educational change in school systems” (Feuer, 2002, p. 8) from the perspectives of field staff and educators in the sites. These descriptions provide a context for the more formal analyses called for by the following two research questions.

The second research question provides the most direct assessment of the efficacy and utility of the model itself. The three-dimensional framework proposed in the model describes an integrated functioning of school structures and personnel “working systemically” in support of student learning. It has been hypothesized that improved systemic work thus leads to improved achievement through its impact on school- and district-level operations and instructional practice. The intent of this research question was to investigate the degree to which sites have been working systemically by analyzing data from two sources, the *Working Systemically* survey and interviews with educators at the various participating schools and districts.

The third research question addresses how student achievement changed in SEDL's sites over time. Student achievement test results are the basis for most state accountability programs across the United States, which link accountability to state curriculum, instruction, and assessment plans. Test results serve as a gauge for whether programs are being successful and therefore provide a referent for whether these components have been implemented effectively.

The relationship between these questions underlies the analytic approach for the report. The first step is to give context to the process of building capacity for systemic work by describing the strategies field staff relied upon to do so. The second step is to investigate whether any longitudinal changes in capacity to work systemically occurred. The third step examines student achievement outcomes among SEDL schools. Finally, the report synthesizes the results to determine whether any relationship between working systemically and student achievement emerged in the sites.

The research design for this report utilized a mixed-method approach that incorporated data from a variety of sources. Therefore, part of the strategy for answering the research questions was to provide a multifaceted perspective on the impact of the *Working Systemically* model in schools and districts. A primary strategy for the resulting analyses was to focus on data and samples that gave the most systematic and complete information to answer our research questions. The research team, therefore, concentrated analyses on the 12 sites that participated in the project through 2005 because these sites enabled us to use consistent datasets while maintaining a sample sufficient for addressing our questions.

For data describing the field staff strategies, the research team used site contact records (SCRs) from August 2003 through 2005. SCR data is described in more detail later in this section. This timeframe was consistent with that used for other sources of data. These 2 years also involved most of the work that had a direct impact on classroom-level instructional practice and so relate most to improving student achievement. Spring 2005 interviews with site educators supplemented the site contact records.

For data regarding systemic work, the study included surveys of site educators from 2003, 2004, and 2005 (conducted in the spring semesters) and fall interviews from 2003 and 2004. These data are described in more detail later in this section. Modifications to the survey instrument between 2003 and 2004 led to structural and item differences. A correlation analysis was conducted to determine whether there was a relationship between the 2003 survey and revised instruments. Analysis results indicated a significant correlation between the 2003 results and subsequent administrations in 2004 and 2005 (Pearson's $r = .77$ and $.73$, respectively; $p < .05$) indicating a relationship between the original and revised instruments.

The correlations suggest that, theoretically, the survey instruments address a similar underlying dimension of systemic work. Therefore, researchers are able to use results from the survey to gauge the relationship between working systemically capacity and achievement for a specific timeframe (e.g., Spring 2003). Item and structural differences preclude the use of the

2003 survey results in a longitudinal analysis with survey data from 2004 and 2005. The survey instruments measure the same construct but in slightly different ways. However, this difference between the instruments is enough to discourage the use of 2003 results in conjunction with 2004 and 2005 data. The situation is akin to administering different math measures in consecutive years. Each set of math scores could be used to investigate general relationships between achievement and other factors (e.g., SES), although the research team would not necessarily be able to determine changes in math achievement from one year to the next if there were differences in the measurement scale.

Researchers collected student achievement data for 2003 and 2004 (and when possible also 2005) and used demographic and achievement indicators to develop matched comparison groups for each grade-level analysis of achievement. Table II.1 presents the full range of data used in analyses for each question.

Table II.1 Project Data by Research Question and Type

Research Question	Spring 2003	Fall 2003	Spring 2004	Fall 2004	Spring 2005
What strategies did SEDL field staff use to build the capacity of low-performing districts and schools to work systemically?		SCRs (N = 12 sites)	SCRs (N = 12 sites)	SCRs (N = 12 sites)	SCRs (N = 12 sites) Interviews (N = 92)
To what extent did low-performing districts and schools increase their capacity to work systemically?	Survey (N = 746)	Interviews (N = 90)	Survey (N = 772)	Interviews (N = 92)	Survey (N = 756)
Did student achievement increase as districts and schools increased their capacity to work systemically?	Student Achievement (N = 22; Match = 22)		Student Achievement (N = 22; Match = 22)		Student Achievement (N = 22; Match = 22)

The next section presents detailed descriptions of sites, data sources, and data samples.

Methods

Sites

Beginning in Fall 2001, SEDL was involved in 23 sites across the five-state region. SEDL began work in three sites during the second quarter of fiscal year 2001. By the third quarter of fiscal year 2002, a total of 16 sites were involved in the project. From school year 2001–2002 to 2004–2005, there were never more than 17 sites working with SEDL at any given time. Over time, several sites dropped out of the project or were discontinued for various reasons while other sites were added as the work moved forward. The reasons for dropping out included lack of commitment, superintendent turnover, and district consolidation. Table II.2 provides a list of all sites along with start and end dates.

Table II.2 Participation of SEDL's Intensive Sites

Site	Start Year	End Year
Oak Hill, AR	2001	2002
Deforest, TX	2001	2002
Piedra Blanca, TX	2001	2002
Piedmont, AR	2001	2003
Cottonwood, NM	2001	2003
Washington City, TX	2001	2003
Southwest City, TX	2001	2003
Delta Village, AR	2001	2004
River City, LA	2001	2004
Mesa, NM	2001	2004
High Meadows, OK	2001	2004
Forked River, AR	2001	2005
Highway Junction, LA	2001	2005
Desert Hills, NM	2001	2005
Farmville, NM	2001	2005
Wrightsville, OK	2001	2005
Bricktown, OK	2001	2005
Athens, LA	2002	2005
Bayou City, LA	2002	2005
Grisham, AR	2003	2005
Roydale, AR	2003	2005
Brookesville, OK	2003	2005
Pineland, TX	2003	2005

As documented in previous reports (Southwest Educational Development Laboratory, 2002, 2004), sites selected for intensive SEDL work had chronic low performance in reading and math and substantial numbers of students who qualified for free and reduced lunch. These characteristics were consistent with the intent of the model to improve student math and/or reading achievement in low-performing districts.

During the duration of the entire 2003–2004 and 2004–2005 school years, SEDL worked with 12 districts (see Table II.3). Eight of the districts were development sites recruited during the first 2 years of the contract. SEDL developed and refined its *Working Systemically* model in these sites. The remaining four sites, recruited during the 2002–2003 school year, served as test sites where field staff could develop systemic work in more structured ways. Thus the work in

these latter sites provided a more formal assessment of the efficacy of the model. The tables in Appendix A present detailed district- and school-level information on each of the 12 sites that were still participating by the end of the project. Four tables outline the eight development sites, presented state-by-state, and a fifth table presents the same categories of information for the four test sites.

Table II.3 SEDL's Intensive Sites 2003–2004 to 2004–2005

State	Development Sites	Test Sites
Arkansas	Forked River	Grisham Roydale
Louisiana	Athens Bayou City Highway Junction	
New Mexico	Desert Hills Farmville	
Oklahoma	Wrightsville Bricktown	Brookesville
Texas		Pineland

Note. N = 12

All sites were low-performing in reading and/or mathematics during the 3 years prior to working with SEDL (see Southwest Educational Development Laboratory, 2002, for more information about selection criteria). Across the 12 sites, 28 schools were involved in the project. In seven of the 12 sites, field staff worked with two schools, while in two sites they were involved in only one. In three other sites, the work spread. In fact, in Farmville, New Mexico, and Brookesville, Oklahoma, the entire district participated (which included five and four schools, respectively), while in Pineland, Texas, field staff worked with the entire preK to fifth grade faculty, bringing systemic work to three campuses. The majority of the sites chose reading (23 of 28 schools, spanning eight of the 12 sites) as their targeted subject area.

The districts varied in size, with total number of schools in each ranging from two to 32 and respective variation in average enrollment (from 695 to 19,356) and number of teachers (from 59 to 1,189). Only two of the 12 sites were in cities (both large cities). The rest of the sites

were divided among a large town (one site), an urban fringe area (one site), small towns (five sites), and rural regions (three sites).

All 28 schools enrolled substantial numbers of students who came from households with incomes low enough to qualify for the federal school lunch program. Only one school did not have a student body where the majority of the population qualified for free or reduced lunch; in 19 of the 28 schools three fourths or more of the students qualified.

The racial/ethnic composition of student populations also varied. Eight of the schools, located in Oklahoma (Wrightsville and Brookesville) and Texas (Pineland), enrolled no racial/ethnic group as a majority but had student populations that were a mix of White, African American, Hispanic, and American Indian students. In seven schools, African Americans predominated; these schools were spread across the Mississippi Delta region in Arkansas (Roydale and Grisham) and in Louisiana (Athens and Highway Junction). White students were the majority in six of the schools, which were found in Forked River, Arkansas, Bayou City, Louisiana, and Bricktown, Oklahoma. The remaining seven schools in which SEDL worked were all in New Mexico; six of these were predominantly Hispanic and one enrolled primarily American Indians. Asian Americans were not represented in high numbers in any of the project's schools, with the highest rate at four percent in a school located in Athens, Louisiana.

Data Sources and Methods

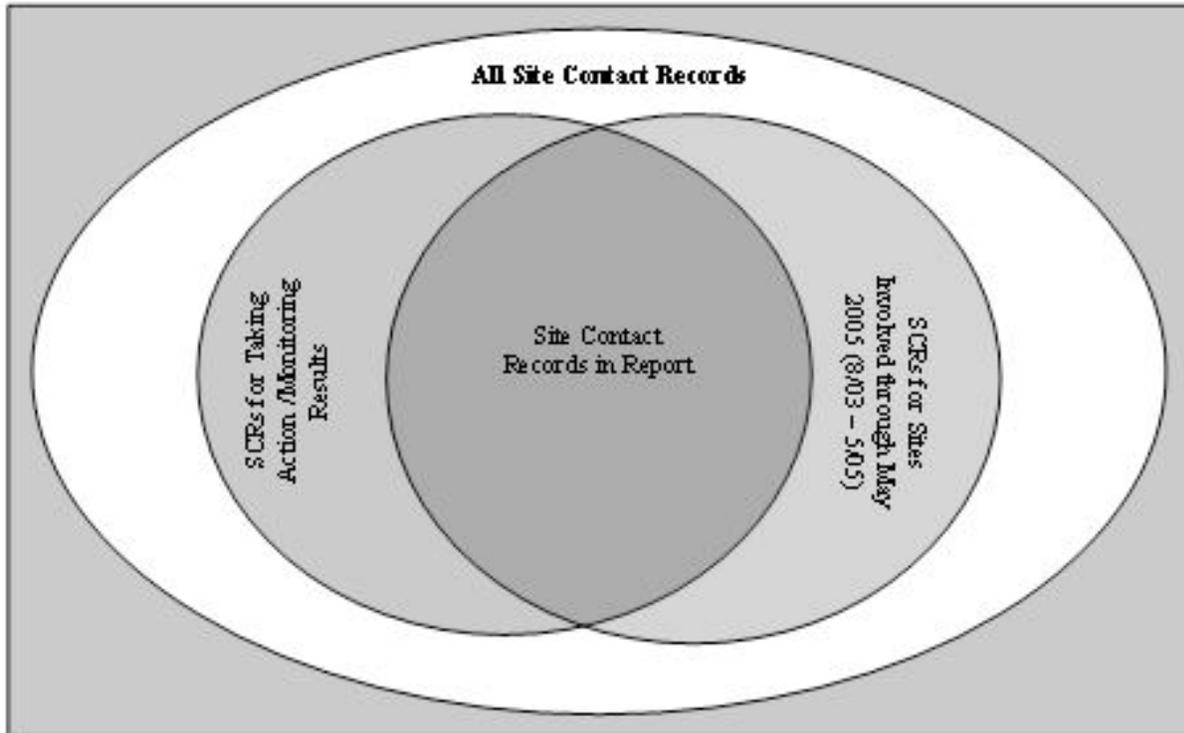
Site Contact Records

Field staff completed electronic site contact records following each interaction with district and school personnel. Most contact was in formal gatherings such as meetings and professional development sessions. The site contact records also contained data pertaining to informal conversations, telephone calls, and written correspondence. The database included prompts to structure the information collected about a particular contact, including narratives of the plan for the particular contact, activities that occurred, the new steps for moving forward, and

the field staff members' reflections about that contact. Each record also listed the location, duration, and other key pieces of information to enable tracking of site contacts and activities over time. Field staff entered this information in a relational database for easy retrieval and regular reviews. A copy of the template used to create site contact records is included in Appendix B.

Given that this is the final report of this project, it was especially important to consider strategies field staff used to develop systemic work that had the most impact on classroom practice. These strategies, in turn, would be the most likely to affect student achievement. This interest led the research team to narrow the examination of the records. One impact of this narrowing was to include only data pertinent to stage 4 of the model, Taking Action and Monitoring Implementation. The other cut, to be consistent with the data sources used for other questions, was to restrict analysis of activity to the last 2 academic years. The intersection of these cuts is represented graphically in Figure II.1.

Figure II.1 SCR Data Sampling Strategy



Field staff marked site contact records as associated with particular levels of the system. Field staff would check off as many levels as they thought were pertinent to the events described in an entry. Research staff checked to see which records were associated with the classroom level because these data, it was expected, would relate to discussions of strategies pertinent to impacting student achievement. In the entire database, less than one percent (34 of 2,302) of the records were tagged as impacting the classroom only. Of these 34, 28 described activity during the last 2 academic years. With respect to the stages of the model, 24 of the 34 occurred during stage 4, Taking Action and Monitoring Implementation.⁶ This pattern was a preliminary index confirming the choice to focus on only those records for stage 4, which were written during the last 2 academic years.

⁶ During earlier stages of the model, such as Analyzing the System, the classroom was on occasion involved during data collection efforts. These records usually do not, however, describe strategies for affecting classroom practice.

Because so few records were tagged as impacting the classroom only, researchers conducted another analysis of the entire SCR database to inform the data sampling strategy. They checked how data for particular levels of the system were distributed over the data associated with different stages of the *Working Systemically* model. To compile counts and proportions for Table II.4, some entries were excluded, such as those that left either the stage of the model or level of the system blank. Entries marked with “state” for level of the system did not describe strategies central to the model and were excluded. Entries marked with more than one stage were also excluded because the purpose of this analysis was to consider which levels of the system were most central to given stages. No site entered stage 5, Assessing and Reflecting on Outcomes, so 14 entries marked erroneously with this stage were also excluded. Due to these exclusions, the total number of entries analyzed is fewer than 2,302, the count for all entries in the database.

When the project first began, there were six stages of the *Working Systemically* model: stage 1, Understanding the System, was previously broken into two stages, Site Entry and Data Scan. For the purposes of constructing Table II.4, entries marked with either of these stages were folded together into those marked with stage 1. The 34 classroom-only entries were recoded as school. Entries marked with district and school were categorized as cross-level. Counts and proportions relative to the resulting total are compiled in Table II.4.

Table II.4 SCR Entries Across the Stages of the *Working Systemically* Model

Stage of model	District only	School only	Cross-level	Total
Understanding the System	70 (41%)	62 (36%)	40 (23%)	172 (12%)
Analyzing the System	70 (36%)	78 (41%)	44 (23%)	192 (13%)
Planning Action	169 (39%)	205 (47%)	58 (13%)	432 (30%)
Taking Action and Monitoring Implementation	146 (30%)	378 (57%)	140 (21%)	664 (45.5%)
Total	455 (31%)	723 (49%)	282 (19%)	1,460 (100%)

Note. Percentages represent proportions of total valid SCR entries.

As Table II.4 shows, there were proportionately more entries associated only with stage 4, Taking Action and Monitoring Implementation (45.5 %, or 664 of 1,460), than with any one of the three earlier stages. Within the count of entries associated with stage 4 ($N = 664$), the fewest number of records were associated with district-level activity only (30%, or 146 of 664), and the greatest number were associated with school-level only (57%, or 378 of 664). Cross-level entries within Taking Action and Monitoring Implementation were proportionately lower compared to other stages of the model. The content of these cross-level fields in stage 4 proved to be relevant to classroom practice because these entries include all the discussions of how leaders monitored teacher activity. It is likely that the cross-level entries in other stages of the model pertained more to collection of data from classroom educators than strategies designed to more directly impact their instructional practice.

Since the records associated with Taking Action and Monitoring Implementation had proportionately more data of interest to this report's analyses, the research team sampled only records associated with this stage. The team did not, however, for the actual sample, exclude records marked with more than one stage so that the full range of activity considered pertinent to Taking Action and Monitoring Implementation would be included in the descriptions of field staff strategy. The sample produced by exporting site-specific text files for all narrative fields and some identifying header information (such as the entry number, date of contact, and type of

contact) is summarized in Table II.5 below. The table also details proportions used from each site's record for this time period.

Table II.5 Fall 2003–Spring 2005 SCR Data Analyzed by Site

Site	SCR Entries Used	All SCR Entries	% SCRs Used	Pages Used
<i>Development Sites</i>				
Athens	31	35	88.6%	88
Bayou City	48	53	90.6%	111
Bricktown	59	68	86.8%	171
Desert Hills	35	48	72.9%	66
Farmville	39	52	73.6%	145
Forked River	70	96	72.9%	187
Highway Junction	53	53	100.0%	161
Wrightsville	32	36	88.9%	57
<i>Subtotal</i>	367	441	83.2%	986
<i>Test Sites</i>				
Brookesville	74	95	77.9%	369
Grisham	23	31	74.2%	40
Pineland	26	36	72.2%	59
Roydale	59	92	64.1%	370
<i>Subtotal</i>	182	254	71.7%	838
Total	549	695	79.0%	1,824

Overall, the sample represents 79% of the data available for these sites during the last 2 academic years. Proportionately, more development entries were sampled (83.2%) than test sites (71.7%) because this cut excluded some of the test sites' planning activity that occurred during this time frame. In contrast, the development sites were further along in the planning stages. Because the authors of two of the test sites' SCRs were more prolific, test site data represents 45.9% of the volume of data within the sample as measured by pages (838 of 1,824 pages).

Interviews

The research team conducted open-ended, one-on-one interviews with educators in Fall 2003 and Fall 2004. Researchers completed 6–9 interviews per site during each round of visits. These interviews consisted of questions about school and district priorities, expectations of teachers and students, alignment, collaboration, resources, and professional development. The interview instrument can be found in Appendix B. Respondents described the actions their schools and districts had taken in each of these areas. The interview questions were the same each year. Researchers both recorded and took notes on interviewees' comments. The research team transformed both sources of data into text files. Table II.6 provides the samples of interviews conducted in each site in the analytic sample. Table II.7 provides a breakdown of interview respondents by role and level.

Table II.6 Fall Interview Sample by Site

Site	Fall 2003	Fall 2004
Athens	8	8
Bayou City	7	7
Bricktown	9	9
Brookesville	7	8
Desert Hills	8	8
Farmville	7	7
Forked River	6	8
Grisham	6	6
Highway Junction	8	9
Pineland	8	8
Roydale	8	6
Wrightsville	5	7
Total	87	91

Table II.7 Sample of Fall 2003–2004 Interview Respondents

Respondents		Fall 2003		Fall 2004	
		N	Percent	N	Percent
<i>District</i>	Superintendents and Assistant Superintendents	17	19.5%	16	18%
	Curriculum Specialists	7	8%	10	11%
	<i>Total District Interviews</i>	24	28%	26	29%
<i>School</i>	Principals and Assistant Principals	24	28%	24	26%
	Teachers	38	44%	30	33%
	Curriculum Specialists	0	0%	9	10%
	Other	1	1%	2	2%
	<i>Total School Interviews</i>	63	72%	65	71%
Total Interviews		87	100%	91	100%
Total Sites		12		12	

In addition to fall interviews, researchers conducted supplementary interviews across sites in Spring 2005.⁷ These interviews focused on school and district practices related to the Professional Teaching and Learning Cycle (PTLC; see Section III for a description of this practice), as well as the long-term prospects for systemic work in the sites. Appendix B contains the interview instrument. Table II.8 reports the sample of educators interviewed in each site.

Table II.8 Spring 2005 Interview Sample by Site

Site	Respondents
Athens	6
Bayou City	9
Bricktown	9
Brookesville	7
Desert Hills	9
Farmville	7
Forked River	8
Grisham	6
Highway Junction	8
Pineland	10
Roydale	7
Wrightsville	6
Total	92

⁷ Interviews were also conducted in Spring 2004 and consisted of questions related to respondents' understandings of alignment. These interviews were analyzed extensively in the Year 4 Report and were not included in this report. For a discussion of these findings, see Southwest Educational Development Laboratory (2004).

A sub-sample of educators from each site participated in the interviews. The sample typically included the superintendent if he/she was involved in the improvement effort, other central office staff who played key roles in the improvement effort, school principals and 2–3 teachers from the schools in which SEDL was working, and content specialists and coaches. Table II.9 describes the roles of Spring 2005 interview respondents and the levels of the system in which they participated.

Table II.9 Sample of Spring 2005 Interview Respondents

Respondents		N	Percent
<i>District</i>	Superintendent and Assistant Superintendent	14	15%
	Curriculum Specialists	8	9%
	Other	4	4%
	<i>Total District Interviews</i>	26	28%
<i>School</i>	Principals and Assistant Principals	21	23%
	Teachers	36	39%
	Other	9	10%
	<i>Total School Interviews</i>	66	72%
Total Interviews		92	100%
Total Sites		12	

The research team developed a quantitative rating scale to evaluate fall interview responses along several dimensions of systemic work. Interview data supplemented some of these quantitative analyses.

Working Systemically Survey

During the contract, the research team developed and refined a survey to measure the degree to which districts and schools were working systemically to maximize student achievement. The original survey consisted of 79 six-point Likert items that assessed respondents' perceptions of whether the district- and school-levels of the system work together to ensure all students can achieve high standards. The items address alignment of standards, curriculum, instruction, and assessment; professional development; the utilization of resources toward school and district priorities; collaboration among stakeholders (school board members,

administrators, teachers, families, and community); and a shared vision that all students can and are expected to achieve high academic standards. Prior to the fourth year of administration in Spring 2004, the instrument was reduced to 50 items. This instrument was also administered in Spring 2005. The *Working Systemically* survey instrument is shown in Appendix B.

The *Working Systemically* survey was administered to the entire faculty at each intensive school site, as well as to all district personnel involved in the district leadership team or working with one of the schools involved with SEDL. Other staff such as teacher aides and district support staff, as well as parents, volunteers, and community representatives, were also included in the samples. Educators completed the surveys during faculty meetings, district leadership team meetings, or school or content team meetings. In a few cases, principals passed out surveys to staff and asked them to return them by a certain time to the office. Table II.10 highlights the sample of survey respondents for sites used in this report.

Table II.10 *Working Systemically* Survey Sample by Site, 2003–2005

Site	Spring 2003	Spring 2004	Spring 2005
Athens	127	49	35
Bayou City	143	133	118
Bricktown	166	112	41
Brookesville	38	75	63
Desert Hills	47	53	51
Farmville	18	45	87
Forked River	39	43	63
Grisham	64	52	47
Highway Junction	77	93	76
Pineland	–	25	101
Roydale	–	61	37
Wrightsville	27	31	37
Total	746	772	756

The survey was administered to all faculty at each SEDL school. Consequently, perspectives on systemic work were largely informed by the opinions of school teachers. Generally, school and district administrators made up the second largest proportion of

respondents. Table II.11 describes the roles of respondents and the levels at which they participated in the sites.

Table II.11 Working Systemically Survey Sample Respondents, 2004–2005

Respondents		2003		2004		2005	
		N	Percent	N	Percent	N	Percent
<i>District</i> <i>t</i>	Administrators	27	4%	19	3%	23	4%
	Other	14	2%	11	1.5%	7	1%
	<i>Total District-Level Surveys</i>	41	6%	30	4%	30	5%
<i>School</i>	Administrators	30	4%	42	6%	45	8%
	Teachers	633	88%	49	84%	480	82%
	Other	18	2.5%	621	7%	29	5%
	<i>Total School-Level Surveys</i>	681	94%	712	96%	554	95%
	Unknown	24	3%	30	4%	172	23%
	Total Respondents	746	100%	772	100%	756	100%
	Total Districts	12	100%	12	100%	12	100%

Previous factor analyses and multivariate analyses of variance (MANOVA) conducted for the Year 4 Report (Southwest Educational Development Laboratory, 2004) indicated that there was significant inter-correlation among the items, making it difficult to identify distinct subgroups of items. This was not surprising considering the survey captured educators' perceptions of the extent to which their schools or districts coordinate processes, resources, and functions to achieve a common objective. Therefore, the analyses for this report used the mean survey score for each respondent as the primary measure of working systemically.

Student Achievement

Student achievement data came from specific state standards-based assessments including the following:

- Arkansas Benchmark Test (ABT)
- Louisiana Educational Assessment Program for the 21st Century (LEAP 21) and the Graduation Examination for the 21st Century (GEE 21)
- New Mexico Standards-based Assessment (NMSA)
- Oklahoma Core Curriculum Test (OCCT)
- Texas Assessment of Knowledge and Skills (TAKS)

The specific testing systems and performance categories are described in more detail in Appendix C. Table II.12 provides a brief description of the data collected for each test by state, including years of data used in the analysis, grades tested, SEDL focus areas tested, and performance categories.

Table II.12 Overview of Student Achievement Data

State	Test	Years Included in Analysis	Analysis Grades	Subjects	Performance Categories
Arkansas	ABT	<ul style="list-style-type: none"> • 2002–2003 • 2003–2004 	<ul style="list-style-type: none"> • 4th grade • 6th grade • 8th grade 	<ul style="list-style-type: none"> • Literacy • Math 	<ul style="list-style-type: none"> • Advanced • Proficient • Basic • Below Basic
Louisiana	LEAP 21 GEE 21	<ul style="list-style-type: none"> • 2002–2003 • 2003–2004 • 2004–2005 	<ul style="list-style-type: none"> • 4th grade • 8th grade • 10th grade (GEE 21) 	<ul style="list-style-type: none"> • English/ Language Arts • Math 	<ul style="list-style-type: none"> • Advanced • Mastery • Basic • Approaching Basic • Unsatisfactory
New Mexico	NMSA	<ul style="list-style-type: none"> • 2002–2003 • 2003–2004 	<ul style="list-style-type: none"> • 4th grade • 8th grade 	<ul style="list-style-type: none"> • Reading 	<ul style="list-style-type: none"> • Advanced • Proficiency • Nearing Proficiency • Beginning Step
Oklahoma	OCCT	<ul style="list-style-type: none"> • 2002–2003 • 2003–2004 	<ul style="list-style-type: none"> • 5th grade • 8th grade 	<ul style="list-style-type: none"> • Reading 	<ul style="list-style-type: none"> • Advanced • Satisfactory • Limited Knowledge • Unsatisfactory
Texas	TAKS	<ul style="list-style-type: none"> • 2002–2003 • 2003–2004 • 2004–2005 	<ul style="list-style-type: none"> • 3rd grade • 4th grade • 5th grade 	<ul style="list-style-type: none"> • Reading 	<ul style="list-style-type: none"> • Met Standard

Note. Inclusion of 2004–2005 achievement data is dependent on availability by state at time of writing.

Table II.12 describes the most complete student achievement data available, which varied across states. Texas and Louisiana released grade-level data relatively quickly, allowing for their inclusion in analyses. While New Mexico also released spring 2005 data, reporting was done at the school level, in many cases combining test results for two grade levels and reporting this as one unit. For example, a middle school with grades 6, 7, and 8 reported the percent meeting the standards for *all* students tested. In New Mexico, this included grades 6 and 8 in 2005. This reporting strategy negated the possibility of including 2005 data in this analysis because grade level achievement was not reported. Table II.13 highlights schools and grade levels utilized for student achievement analyses.

Table II.13 Schools Included in Student Achievement Analysis

State	School	Tested Grade(s)	Tested Subject	SEDL Academic Focus
<i>Arkansas</i>	Grisham–Elementary School C	4	Math	Math
	Grisham–Middle School A	6	Math	Math
	Forked River–Elementary School B	4, 6	Literacy	Reading
	Forked River–High School B	8	Literacy	Reading
<i>Louisiana</i>	Athens–Middle School A	8	Math	Math
	Bayou City–Middle School B	8	Math	Math
	Bayou City–High School A	8	Math	Math
	Highway Junction–Elementary School A	4	English/Language Arts	Reading
	Highway Junction–Elementary School B	4	English/Language Arts	Reading
<i>New Mexico</i>	Desert Hills–Elementary School A	4	Reading	Reading
	Desert Hills–Middle School A	8	Reading	Reading
	Farmville–Elementary School B	4	Reading	Reading
	Farmville–Elementary School C	4	Reading	Reading
	Farmville–Middle School C	8	Reading	Reading
<i>Oklahoma</i>	Bricktown–Middle School A	8	Reading	Reading
	Brookesville–Elementary School D	5	Reading	Reading
	Brookesville–Middle School B	8	Reading	Reading
	Wrightsville–Elementary School A	8	Reading	Reading
<i>Texas</i>	Pineland–Elementary School A	3, 4, 5	Reading	Reading

Note. The unit of analysis for student achievement data is the grade level. In other words, there are 22 grade-level analyses across the 19 schools in the 11 sites presented in this table.

Strengths and Limitations of Data

The types of data collected for this report have several strengths. Complex social phenomena, such as the systemic inter-workings of schools and districts, require multiple methods of data collection and analysis to best capture and make inferences about these complexities (Greene & Caracelli, 1997). To this end, collecting qualitative and quantitative provided both depth and breadth (Teddlie & Tashakkori, 2003) in the examination of the impact of SEDL's *Working Systemically* model.

Qualitative data obtained from site contact records documented all contacts with districts and schools and gave the research team a rich history of how systemic work was built in each site. The interviews provided additional qualitative and quantitative perspectives from the administrators and faculty in these districts and schools. Data from the *Working Systemically* survey allowed for quantitative analyses of educators' perceptions of district- and school-level practices over time. Thus, analyses of these data permitted both convergent and complementary triangulation of emergent findings (Erzberger & Kelle, 2003).

This approach also had limitations. Site contact records, the *Working Systemically* survey, and interview data were subjective in nature and may have not always conveyed accurate representations of what was occurring in the sites.

As discussed previously, student test data were collected for the purpose of examining achievement patterns across the various SEDL sites. Ideally, student achievement trends would be examined using individual student-level data. This type of data can be used to conduct more in-depth analyses due to the level of precision they provide. Unfortunately, federal regulations surrounding individual student data prevented their use due to concerns about confidentiality. Thus, SEDL conducted analyses of the best available test data from each state for insights they provided regarding the implicit goal of increasing student achievement.

More generally, analysis of student data had to acknowledge that state definitions of proficiency changed as the accountability systems matured. In some instances, states changed the tests themselves, thus restricting certain longitudinal analyses. For example, Texas switched tests from the Texas Assessment of Academic Skills (TAAS) to the Texas Assessment of Knowledge and Skills (TAKS) in 2002–2003.

Another limitation stemmed from the fact that, currently, most state accountability systems emphasize cross-sectional grade-level assessment. For instance, Arkansas administered its benchmark tests at grades 4, 6, and 8. Thus, analyses of student outcomes necessarily relied on year-to-year comparisons of grade-level data (e.g., 2003 grade 4, 2004 grade 4, 2005 grade 4). The most apparent consequence of this strategy was that even within a single school, the population of students in one grade may have been substantially different from one year to the next, potentially skewing achievement results.

In addition to limitations outlined above, modifications to the *Working Systemically* survey led to substantial differences between the 2003 and 2004 instruments.⁸ In addition, interview practice and protocols also varied at the beginning of the project, leading to inconsistent data. Although some variability is inevitable when using interview strategies, interview procedures and conventions were more structured in Fall 2003 and 2004, yielding a more standard format.

Finally, attrition played a role in the overall structure of the available data. For instance, four sites (Deforest, Piedmont, Oak Hill, and Piedra Blanca) discontinued their work with SEDL prior to collection of survey data in spring 2003. Conversely, Roydale and Pineland did not join SEDL as test sites until early 2003.

⁸ Changes included a different response scale and fewer items.

The research design for this report attempts to address some of these limitations through the judicious use of the most consistent data available for the project.⁹ This approach underlies the sampling frame for each research question. In light of this strategy, the following sections highlight the analytic plans for each research question.

Analyses

This section describes the analytic approaches used for each of the three research questions. Findings from these analyses are presented in sections III, IV, and V, respectively.

What Strategies Did SEDL Field Staff Use to Build the Capacity of Low-Performing Districts and Schools to Work Systemically?

The analytic approach to this research question is a composite based on actual field staff strategies. It is similar to what Yin (1984) calls a “multiple-case report,” wherein the researcher “draws appropriate examples” from cases but no one “is presented as a single-case study” (p. 130). The multiple-case report presents information from all cases organized around topics.

Research staff used drafts of the field staff’s *Working Systemically* toolkit¹⁰ as a way to identify domains¹¹ of field staff strategies. These documents describe what a facilitator should do during each of the model’s five stages: Understanding the System; Analyzing the System; Planning Action; Taking Action and Monitoring Implementation; and Assessing and Reflecting on Outcomes. Researchers also reviewed documents available that they and the field staff had generated internally during the project. Examples of these are notes taken during project meetings and tools the field staff used with educators in their work.

⁹ For a description of complete SEDL databases, refer to Appendix E.

¹⁰ The *Working Systemically* toolkit is based on extensive revision of the project protocols. Because this research study was also a development project, the protocols were revised periodically during the course of implementation to integrate what field staff were learning in the sites about how to meet educators’ needs better. The toolkit represents, in effect, the latest version of the model’s protocols. Research staff read through each draft as they were developed during the course of the project.

¹¹ LeCompte & Schensul (1999) offer an overview of domain analysis. Developed by James Spradley, domains are classes of items, such as objects, things, ideas. The domains are the roots of a taxonomy that organizes conceptually the relationships among the items.

The domains represent core strategies that summarize broad approaches to developing systemic work, which are operationalized through more specific actions. These specific strategies associated with the core strategy domains were identified using a combination of methods. First, researchers examined the structure of the site contact record database. Whenever possible, researchers conducted analyses using the fields where field staff made selections from standardized check boxes included in the site contact record database. To explore a given category of field staff action further, research staff then performed more targeted searches of the database. They were able to review information both by examining the more quantitative data (such as number and type of participants) as well as reading through the longer portions of narrative text embedded in the designated fields.

The content, style, and amount of writing in the fields into which text was entered varied. As a result, for systematic review of data reported only in these textual fields, the data were exported into site-specific text files, producing approximately 1,800 pages of text. These data were imported into a qualitative software package and coded for descriptions of specific strategies. Codes were also developed to organize data that were not describing strategies themselves but other related background information, such as assessments of how the work was going. The technique for coding relied primarily on iterative text searches for common terms related to the strategy of interest, as well as reading and reviewing extensive passages surrounding relevant “hits” from text searches. Topically coded reports corresponding to a given strategy or related subject, the outcome of this analytic process, were then reviewed for general patterns in the data. Subsequent search and review added more data to these reports until a cross-site analysis was complete.

Where SCR data were missing regarding a particular strategy, research staff triangulated analyses using Spring 2005 exit interviews ($N = 92$; see Appendix B for a copy of the instrument). Relevant individual members of the field staff were also consulted to clarify questions and supply detail where necessary.

The Spring 2005 exit interviews offered the educators' perspectives, helping researchers cross-check field staff's interpretations of how the work was progressing. Educators answered questions regarding the history of the project and predictions for its sustainability. These data were also used for additional detail about field staff strategies, particularly with respect to the Professional Teaching and Learning Cycle (PTLC), the primary method of directly improving classroom instruction. Researchers entered all data from the interviews into the qualitative software package. Each interview was also coded by question, for site, and for role of the respondent within the levels of the educational system. Researchers could then easily review all answers to a particular question, broken down, if desired, by the demographic variables.

When review of the site contact records, spring interviews, and project documents did not suffice to complete a given analysis, researchers held peer debriefings directly with pertinent field staff members. These were conducted most often to gather additional information about the sites' action plans and regarding PTLC activities and how they were monitored.

For both site contact records and interviews, when deciding which text should be included for purposes of illustration in this report, research staff selected text that

- illustrated typical field staff practice;
- represented activity across a variety of sites; and
- provided relatively succinct summaries of activity that illustrated the strategy of interest.

Research staff represented cuts in textual data with ellipses in parentheses [(. .)], so as to distinguish them from use of ellipses in the original. Researchers corrected spelling and typographical errors with no notification to the reader but otherwise bracketed additions, deletions, and substitutions. Identifying information was deleted, with more generic terms substituted (e.g., changing Mr. Name to [Educator]). Personal pronouns referring to individual field staff members were not used to protect their anonymity.

To What Extent Did Low-Performing Districts and Schools Increase Their Capacity to Work Systemically?

Fall interview and *Working Systemically* survey data were used to determine whether sites developed practices and goals related to working systemically over time. Quantitative analyses focused on the issue of change over time, while interview data elaborated on events in sites with notable results and suggested possible explanations for differences in results between measures in those sites.

Fall Interview Ratings

Interviews were rated on four outcomes pertinent to the concept of working systemically. Each area represented practices that incorporated one or more of the competencies defined in the *Working Systemically* model. The four areas were:

- **Expectations and Priorities:** Expectations of teachers and school/district priorities are both focused on improving student learning in the SEDL focus subject area.
- **Alignment:** There is alignment between curriculum and the state standards in the SEDL focus subject area.
- **Collaboration:** Educators collaborate to improve student learning in the SEDL focus subject area outside of professional development sessions.
- **Professional Development:** Educators use school- or district-sponsored professional development to improve student learning in the SEDL focus subject area.

These outcomes also related directly to interview topics: school and district priorities, expectations of teachers and students, alignment of curriculum and instruction to standards, collaboration, availability of instructional resources, and professional development (the interview instrument is included in Appendix B). Rubrics for each outcome were developed to assign ratings to each interview transcript using a scale of 1 (lowest) to 5 (highest).¹²

¹² Data on instructional resources did not go into enough depth to be useful, so a rubric was not developed for this subject.

Although the rating scales mirrored the questions in the fall interview instrument, respondents often discussed a particular topic at various points in the interview. For example, a respondent might refer to alignment of curriculum in responding not only to questions on alignment, but also to questions on collaboration or professional development. To account for these patterns of response, ratings derived from a holistic scoring process that utilized the entire interview transcript to assign ratings for each scale, similar to processes often used to rate essay portions of standardized assessments (such as the Graduate Record Examination).

The rubrics measured the extent to which practices and structures related to each outcome had become formally integrated and ritualized into the system of the district. In other words, the rubrics characterized activities that spanned the components and levels of the district as described in the model. Each scale is described briefly below. The rubrics themselves are presented in Appendix D.

Ratings for expectations and priorities focused on the centrality of student learning or student achievement in discussions about goals and priorities and the congruence between these stated goals and expectations and requirements of teachers. Transcripts receiving ratings of 1 contained data indicating student learning or achievement was not a high priority in the site (as opposed to other priorities, such as safety or discipline, for example). A rating of 5 indicated that student learning or achievement was a high priority and that teachers and administrators throughout the district were expected to engage in certain practices related to this goal. One respondent, for example, discussed districtwide requirements that teachers administer pre-assessments to students to individualize instruction and that principals check teachers' lesson plans for adherence to standards on a weekly basis.

For the rubric on alignment, the presence of a written curriculum aligned to state standards was a focal point. A rating of 1 on this rubric indicated that a school or district had no written curriculum or that the curriculum was not aligned to standards. A written curriculum was defined for this purpose as a district's or school's interpretation of state standards into topics

and/or lessons on a particular subject to be taught within a grade level. Reports of use of state standards, even if those standards were highly specified by subject and grade level, was not sufficient to constitute a district or school curriculum because it was unclear how systematically those standards guided the enacted curriculum. A rating of 5 depicted alignment between a written curriculum and state standards, with the curriculum spanning grade levels and being used throughout the district. Such a curriculum might also have a scope and sequence specifying particular units to be taught throughout the year.

The collaboration rubric measured sites' use of formal (i.e., endorsed) collaboration among educators for purposes other than professional development, particularly planning curriculum and instruction. Formal collaboration was defined as scheduled meetings that school or district leaders sponsored or endorsed. Examples included grade-level meetings, department meetings, horizontal or vertical team meetings, and principals' meetings. Ratings depicted the formality, frequency, and scope of such meetings throughout the site. A rating of 1 indicated reports of little or no collaboration or that respondents viewed collaboration negatively. A rating of 5 meant formal collaboration took place regularly among teachers and administrators throughout the district.

Regularity and focus were also key issues for the rubric on professional development. A rating of 1 indicated professional development activities were not related to student learning needs in the site (as opposed to other needs such as state mandates unrelated to curriculum or instruction, teachers' individual interests, or in one case, "mundane housekeeping chores"). A rating of 5 suggested that professional development in the site was focused on building skills related to student learning needs and provided for consistent follow-up or monitoring activities to ensure understanding and implementation.

Two research team members assigned ratings. Inter-rater reliability was established by Cronbach's alpha ($\alpha = .638$). The research team defined agreement as two raters differing by only one point or less when rating the same interview.

Ratings were then summed across rubrics to compute sum scores for each interview. The minimum possible sum score for each interview was 4, which corresponds to a rating of 1 on each scale. The maximum possible score was 20, which corresponds to a rating of 5 on each scale. Aggregating individual sum scores yielded mean site sum scores for both 2003 and 2004. Mean site scores were also computed for each rubric for both 2003 and 2004. These scores represented mean ratings across individual interviews using the scale of 1 to 5 rather than sum scores because the sample sizes for each site in each year varied. Mean site sum scores served as the basis for comparing sites descriptively. Analyses of variance (ANOVAs) used individual sum scores to compare means for each site over time and to compare means of various groupings of sites and respondents over time.

Working Systemically Survey

The *Working Systemically* survey, consisting of 50 Likert items, lent itself readily to quantitative analysis. Reliability for the survey using Cronbach's alpha was high ($\alpha = .97$). The research team assigned a mean score to each survey reflecting the Likert scale of 1 to 6 (1 = "Never True," 6 = "Always True"; see Appendix B for details on other response categories), aggregated these by site, and then assigned a mean site score using the same scale for each year the survey was administered. Site mean scores were used in descriptive tables and graphs. Individual respondent mean scores were used in univariate analyses of variance (ANOVAs) to compare means for each site over time and to compare means of various groupings of sites and respondents over time. Sample sizes for all analyses conducted are provided in Section IV. As discussed earlier in this section, inconsistencies in the survey instrument between 2003 and 2004

prevented 2003 data from being included in longitudinal comparisons. However, descriptive data for 2003 for each site are presented for examination in Section IV.

Fall interview sum scores and *Working Systemically* survey mean scores were highly correlated (Pearson's $r = .64$), adding to confidence in their validity. Despite this high correlation, the measures were both quantitatively and qualitatively different instruments and were thus appropriate to independent analysis.

Did Student Achievement Increase as Districts and Schools Increased Their Capacity to Work Systemically?

The analytic approach to this question encompassed two dimensions. First, researchers analyzed student data for tested grades to determine whether there were any consistent patterns of achievement among SEDL schools. As mentioned previously, student achievement data for this report were limited to the proportion of students meeting performance levels as defined by state assessments. Consequently, only non-parametric techniques were suitable for assessing differences among proportions. The basic question underlying these analyses was whether SEDL schools exhibited significant changes in the proportion of students meeting state standards over time. In addition to year-to-year changes, analyses also incorporated the use of a composite comparison group as a point of reference for SEDL school achievement.

The second dimension of the analytic approach examined the relationship between measures of systemic work and student achievement. Results from the *Working Systemically* survey and interviews were examined together with achievement data to determine the degree of association between these measures.

In order to bound the achievement data in the range of the *Working Systemically* survey results and interview ratings, most of the analyses for this question focused on Spring 2003 and Spring 2004 test administrations. When available, spring 2005 data were also included.

The primary analytic approach for the analyses of student achievement data was to apply the chi-square test of association to the number of students meeting state-defined proficiency levels from one year to the next.¹³ The basic question posed by the test is whether the differences among proportions are attributable to sampling error (Glass & Hopkins, 1996). In other words, when we see increases or decreases in proportions from one year to the next, was that change more than you would expect beyond chance?

The chi-square analyses for this report were applied to the proportion of students assigned to specific performance-level categories across the SEDL sites. The data for each grade-level analysis were organized as contingency tables. Table II.16 illustrates a sample 2 x 2 contingency table for a fictional analysis.

Table II.14 Sample Chi-Square Table for Third-Grade Reading in SEDL School A

	2002–2003	2003–2004	Total
Met Standard	22 (44%)	28 (56%)	50 (50%)
Did Not Meet Standard	28 (56%)	22 (44%)	50 (50%)
Total	50	50	100

In this hypothetical data set, although there was a slight increase in the number of students meeting the standard in 2003–2004 relative to 2002–2003, the chi-square analysis suggests that the change was not significant: $\chi^2 (1, N = 100) = 1.44, p > .05$.

Analyses concentrated on the tested grade levels in each SEDL school listed in Table II.13. In most sites, SEDL's work happened in several schools, usually at least one elementary school and one secondary school. However, this pattern was not consistent. For example, in one site field staff worked with only one elementary school. Given potential differences in the number of schools that may have represented a tested content area in a site, aggregation became

¹³ While a chi-square goodness-of-fit test can be used to determine whether observed proportions differ from theoretical proportions, the chi-square test of association uses available data to estimate the expected proportion.

problematic. Even within a state, sites were not always working on the same content area. The work in Louisiana, for instance, included two math sites and one reading site representing a total of five schools.

Although achievement data were available for school years prior to 2002–2003 for some sites, the report focuses on reported results for 2002–2003 through 2004–2005. The rationale for this approach is that 2002–2003 represents the first achievement data point that is consistent across states. For example, while Louisiana reports data for years prior to the beginning of SEDL's work, 2002–2003 was the first administration year for the New Mexico Standards-based Assessment and the new Texas Assessment of Knowledge and Skills. The use of data for these years (2002–2003 to 2004–2005) was also more consistent with the timeframe for other data sources utilized in this report (e.g., *Working Systemically* survey and interviews).

For purpose of comparison, SEDL schools were also matched to similar schools on demographic characteristics and prior achievement. The matching variables were consistent with previous research (Borman & Hewes, 2001; Datnow et al., 2000; Laczko-Kerr & Berliner, 2002). The research team selected a group of similar schools based on the following school-level variables: enrollment, percent minority, percent of students on free and reduced lunch, percent English language learners, and prior achievement. The mean number of students meeting the standards was then calculated for each matched set providing a comparison group for each SEDL school. This strategy provided a more balanced comparison because a single school may experience unique situations that produce peculiar outcomes.¹⁴

Analysis of achievement data for composite schools was consistent with that of SEDL schools. The results from these analyses helped determine whether achievement patterns for students differed between SEDL sites and non-SEDL sites with similar characteristics.

¹⁴ There were a few cases where a comparison group or single school could not be identified that matched with the SEDL school on initial achievement. In these cases, researchers created a “best match” using schools selected on demographic criteria.

Differences between these entities provided at least some insight regarding the relative impact of the *Working Systemically* model. For example, did SEDL schools exhibit larger and significant increases in the proportion of students passing a state test from 2002–2003 to 2003–2004 relative to non-SEDL schools?

In addition to grade-level analyses of SEDL schools compared to matched schools, the research team also attempted to provide a more global examination of the relationship between achievement and systemic work in sites. Several correlation coefficients were calculated between survey results, interview ratings, and achievement results to investigate this relationship.

In order to utilize student achievement across grade levels, subject areas, and states, the achievement variable was represented as the proportion of students meeting standards in reading or math (0.0–1.0). The cases reflected the grade-level analyses. Mean total scores for the *Working Systemically* survey and interview ratings were calculated for each site. Site-level means for the survey and interview ratings were then assigned to the corresponding grade-level units in each site to calculate Pearson-*r*.

III. FIELD STAFF STRATEGIES

What Strategies Did SEDL Field Staff Use to Build the Capacity of Low-Performing Districts and Schools to Work Systemically?

This section describes the specific strategies underlying the *Working Systemically* model that were thought to be most related to classroom practice and improving student achievement. The field staff's site contact records pertaining to stage 4 of the *Working Systemically* model, Taking Action and Monitoring Implementation, provide much of the data for the discussion.

During earlier stages of the work, field staff and site leaders drafted plans outlining processes to bring about the alignment of the components, especially curriculum, instruction, and assessment with the state standards. Once the work reached stage 4, Taking Action and Monitoring Implementation, field staff became more intensively involved in building capacity at all levels of these low-performing educational systems. In this stage, the work integrated classroom-level educators' activities, with leaders ideally participating in them or at least monitoring their progress. In other words, as the action plans were enacted, the project focused even more on the competency of "creating coherence" across all levels of the system. The work in this stage put great emphasis on bringing the competencies of systemic work to the classroom. Specifically, a greater number of teachers became integrated into leaders' efforts to create coherence, use data, ensure continuous professional learning, and build relationships. District and school leaders, when necessary, responded to changing conditions and adjusted plans for improvement.

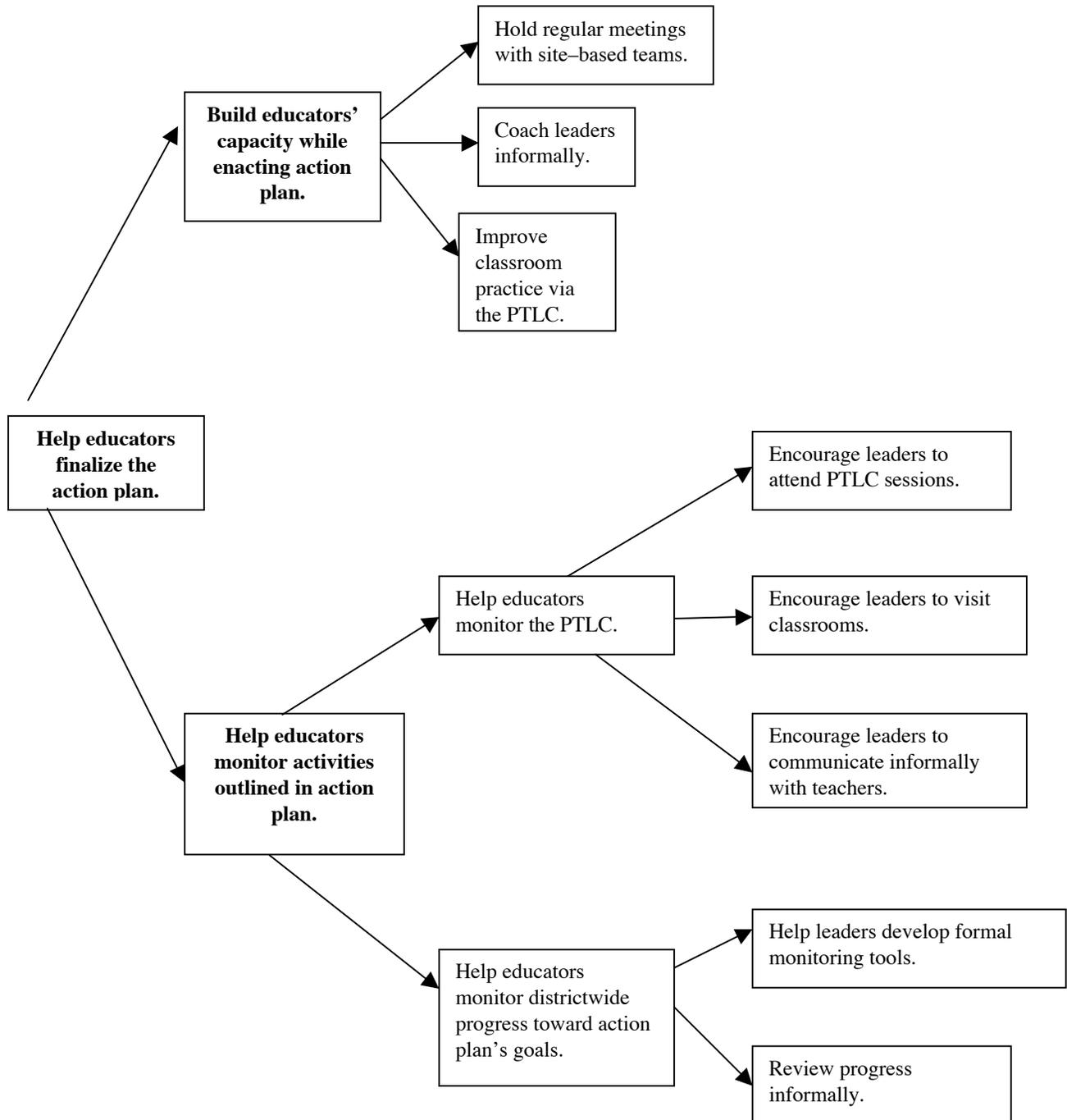
Findings

The Core Strategies

The following section describes the strategies field staff used during this stage of the model to develop systemic work. Figure III.1 outlines graphically the connections among them. The three leftmost boxes with bolded text represent three domains of field staff strategy,

corresponding to the three major analyses presented in Section III. These three “core strategies” can be thought of as general approaches to developing systemic work, which must be enacted via more specific and concrete strategies. In other words, an educational system cannot just “build capacity”; capacity is built via several activities. Moving right, other text boxes represent the specific strategies related to these three “core strategy” domains, which are described in more detail in the subsections to follow. The figure does not provide a complete outline of Section III. For example, some material to be discussed does not describe strategies but rather provides background material or examines issues pertinent to them.

Figure III.1 Field Staff Strategies During Taking Action and Monitoring Implementation



Starting from the left in Figure III.1, one core strategy during stage 4 was to finalize the action plans. Field staff reviewed the content of drafts of the action plans with the district leadership team members. To promote alignment of classroom instruction with curriculum, assessment, and the state standards, all field staff included the Professional Teaching and Learning Cycle (PTLC) in the action plans; it is described in detail below. They also tried to ensure that other initiatives with which the site was engaged were aligned with the goals of the action plans. Once the action plans were finalized, field staff supported educators with two other core strategies (see bolded text boxes connected by arrows to the leftmost box). Field staff worked to build educators' capacity to do the work outlined in the plan. They also helped leaders monitor both the PTLC and districtwide progress toward goals set out in the action plan.

A second core strategy was to build educators' capacity while they enacted the action plan. There were three primary strategies through which field staff sought to build capacity, represented in the upper right of Figure III.1 by the next three boxes. One strategy was to create district leadership teams and, in some sites, teams of administrators or school-level teams to facilitate collaborative decision making. Secondly, field staff coached leaders informally. These two strategies focused especially on leadership. As indicated by the third text box in this grouping, field staff relied primarily on the PTLC to focus on teachers and coaches, including principals and other leaders as well. These efforts concentrated on content and pedagogy at the school level during collaborative professional development sessions with educators.

A third core strategy during this stage in the work was to help educators monitor the progress toward goals set out in the action plan. Associated strategies are represented in the lower half of Figure III.1, starting with the first box connected to the leftmost one. Field staff strengthened leaders' ability to monitor progress in two areas (the next two boxes, moving right in the figure): PTLC activities and progress toward districtwide goals. There were three primary strategies field staff used to monitor the PTLC. One was to ask leaders to attend teachers' PTLC collaborative work sessions. Another strategy was to train leaders how to do classroom visits for

the purpose of monitoring how teachers applied what they were learning during their PTLC meetings, using a procedure often referred to as “classroom walkthroughs.” Additionally, field staff encouraged leaders to communicate informally with teachers, teaching them a technique called “one-legged” interviews, the referent being the casual stance of leaning on one leg during these conversations. Field staff took two approaches to monitoring progress toward districtwide goals. One was to use formal tools developed for that purpose. The other was to use some of the district leadership meeting time to conduct informal reviews of activity related to the action plans and their goals.

After examining these three core strategies and associated strategic activities, Section III concludes with an overview discussion. It includes summary comments based on insights field staff and educators had about the advantages and challenges of the field staff's approaches to developing systemic work. It also discusses issues related to sustainability.

Finalizing the Action Plans

A major strategy field staff used to provide a framework for their systemic reform activities was to develop an action plan for the district work. Together with members of the district leadership team, they used the template in Figure III.2 to help them define what parties and activities would be within the scope of action, when tasks would be completed, and what would be considered evidence of accomplishment. In stage 4, this plan was usually still being finalized.

Figure III.2 Action Plan Template

Description of the Ideal State: _____
 Critical Element: _____

Action Steps to Be Taken	Person(s) Responsible for Accomplishing Action	Date Action Will Be Completed	Evidence of Accomplishment

Members of the district leadership team, based on the analyses of their educational system completed in earlier stages of the model, had identified a root cause of their district's chief problems. Then they had devised a "Description of the Ideal State": a situation in which the root cause of the problem was addressed. "Critical elements" of that ideal state became the foci of the action plans. In most sites, site coordinators drafted the action plan. District leadership team members discussed and made changes to it.

The most common areas of focus for site action plans are outlined below in Table III.1. For the purposes of building the table, researchers considered "alignment" to be addressed not only when it was explicitly named, but also when it was implied. For example, the plan for Desert Hills includes "assessment informs instruction" and "curriculum and instruction based on standards" as areas of focus. Since these phrases describe alignment, that site was included in the "alignment" column of Table III.1.

Table III.1 Most Common Areas of Focus for Site Action Plans, 2004–2005

Site	Alignment	Curriculum	Instruction	Assessment	Standards
<i>Development Sites</i>					
Athens	X	X	X	X	X
Bayou City	X	X	X	X	X
Bricktown	X	X		X	
Desert Hills ^a	X	X	X	X	X
Farmville				X	X
Forked River	X	X	X	X	X
Highway Junction	X	X	X	X	X
Wrightsville	X		X	X	X
Subtotals	7	6	6	8	7
<i>Test Sites</i>					
Brooksville	X	X	X	X	X
Grisham	X		X	X	X
Pineland	X	X		X	
Roydale	X	X	X	X	X
Subtotals	4	3	3	4	3
TOTALS	11	9	9	12	10

Note. Plans from this academic year were sampled because some of the sites did not have action plans during the prior year.

^a Desert Hills had two school action plans rather than one for the district. These plans had the same areas of focus.

A few action plans addressed other areas of focus, especially those produced by sites that joined the project in its earlier years. For example, only Highway Junction included family and community involvement as a critical element, and both Forked River and Desert Hills mentioned collaboration. Bricktown’s district leadership team put in communication of the action plan as one of its critical elements as well as professional development for school staff related to literacy, its curricular area of focus. Farmville’s action plan had implementation of improvement efforts as an area of focus. All of these elements were consistent with the *Working Systemically* model. In later years, however, field staff activity targeted components represented in the *Working Systemically* model that were considered more directly related to student achievement: curriculum, instruction, assessment, and the standards. Assessment was mentioned in all of the

plans, perhaps because of the recent emphasis on student achievement in districts' accountability requirements. It is also important to point out that Table III.1 summarizes only the named areas of focus. The tasks associated with these main areas may include activities related to other components, such as instruction and curriculum, which were mentioned slightly less often than assessment and standards. By academic year 2004–2005, all of the sites had goals compatible with work toward the alignment of the components included in Table III.1.

A comparison of the time spent on developing the action plans and other earlier work showed a difference between development and test sites. Due to the project having more developed protocols in its later years, test sites proceeded more quickly through this process of entering sites, obtaining site commitment, forming the district leadership team, and doing the work necessary to produce an action plan informed by data collection and analysis.

Table III.2 Time Spent Prior to Stage 4, Taking Action and Monitoring Implementation

Site	Months Spent Prior to Stage 4
<i>Development Sites</i>	
Athens, LA	13
Bayou City, LA	13
Bricktown, OK	12
Desert Hills, NM	13
Farmville, NM	15
Forked River, AR	19
Highway Junction, LA	11
Wrightsville, OK	13
	Mean = 13.6 months
<i>Test Sites</i>	
Brookesville, OK	9
Grisham, AR	8
Pineland, TX	13
Roydale, AR	9
	Mean = 9.75 months

The Professional Teaching and Learning Cycle (PTLC)¹⁵

According to Mujis, Harris, Chapman, Stoll, and Russ (2004), disadvantaged schools benefit from concentrating specifically on the goal of improving student achievement. With the push for improvement in student performance central to most state accountability plans as required by the federal government under the No Child Left Behind Act of 2001, participants in the project understood even more how important it was to direct districts' and schools' attention squarely on this topic (Elmore et al., 1996; Odden, 1998). Alignment of curriculum, instruction, and assessment with the state standards at the school and classroom level can improve student achievement (Newman et al., 2001). SEDL introduced the Professional Teaching and Learning Cycle (PTLC) in all of the sites by January 2004 as a way of catalyzing this alignment process. After this date, the PTLC was the principal professional development strategy field staff used to build teacher capacity. Since most of the content of action plans written prior to this date was compatible given the shared emphasis on these components, the PTLC represented a way to standardize field staff's approaches to their alignment. Activities related to teaching educators how to use the PTLC were a part of all action plans as they were completed or updated for the 2004–2005 academic year.

The PTLC organized teachers' collaboration in ways that helped align classroom instruction to the state standards. It became an integral part of SEDL's systemic work during this project, focusing on all five competencies. As a job-embedded, collaborative process, it offered continuous professional learning to participants. They strengthened content and pedagogical skills as they deepened their professional relationships. It was the means by which the field staff ensured that their earlier efforts among leaders to promote the importance of aligning the components would have a direct impact on the classroom. The PTLC was designed to be flexible

¹⁵ The Professional Teaching and Learning Cycle was adapted from the Professional Teaching Model of the Charles A. Dana Center.

enough to work with either the math or reading sites. Figure III.3 details central features of the PTLC. This figure is the actual tool written and used by field staff in the sites.

Figure III.3 Professional Teaching and Learning Cycle (PTLC)



Prior to beginning the first Professional Teaching and Learning Cycle (PTLC), collaborative planning teams (grade level, vertical, and/or departmental teams) examine student achievement data and a scope and sequence of their grade level/course curriculum. Teams use these data to guide the decision of which standard(s) to address first.

Phase I—Study:

Teachers work in collaborative planning teams to critically examine and discuss the learning expectations from the selected state standards.

Teams develop a common understanding of:

- The concepts and skills students need to know and be able to do to meet the expectations in the standards
- How the standard fits within a scope and sequence of the district curriculum
- How the standards for a grade/course are assessed on state and local tests

Phase II—Select:

Collaborative planning teams research and select instructional strategies and resources for enhancing learning as described in the standards.

Teams:

- Identify effective research-based strategies that could be used to support learning in the selected state standards
- Select the most-appropriate resources to use with the strategies
- Determine how often to use a strategy, or strategies, and for what purpose
- Agree on appropriate assessment techniques that could be used to provide evidence of student learning

Phase III—Plan*:

Collaborative planning teams work together to formally plan a lesson incorporating the selected strategies and come to agreement on the type of student work each teacher will bring to the analysis phase that reveals evidence of student learning.

Teams:

- Develop a common formal plan outlining the lesson objectives (relevant to the standards), the materials being used, the procedures, the time frame for the lesson, and the activities in which students will be engaged
- Discuss how the lesson can be modified to address the diverse learning needs of students
- Decide what evidence of student learning will be collected during the implementation

* *The Study, Select, and Plan phases are dynamically linked and iterative to ensure alignment of instruction and assessment with the standards. When teachers are in the plan phase, their conversations should refer to the standards and instructional strategies again (Study and Select) to guarantee that the lesson is aligned with the standards.*

Phase IV—Implement:

Teachers teach the planned lesson, make note of implementation successes and challenges, and gather the agreed upon evidence of student learning.

Teachers:

- Deliver the lesson as planned in the specified time line
 - Record implementation results, especially noting where students struggled and/or where instruction did not achieve expected outcomes
 - Collect the agreed upon evidence of student learning to take back to the collaborative planning team
-

Phase V—Analyze*:

Teachers gather again in collaborative teams to examine student work and discuss student understanding of the standards.

Teams:

- Revisit and familiarize themselves with the standards prior to analyzing student work
 - Decide on the process (protocol) for examining and analyzing student work
 - Analyze sampling of student work for evidence of student learning
 - Discuss whether or not students have met the expectations outlined in the standards and make inferences about the strengths, weaknesses, and implications of instruction
 - Identify what students know and what needs to be strengthened in future lessons
-

Phase VI—Adjust*:

Collaborative teams reflect on the implications of the analysis of student work. Teachers discuss alternative instructional strategies or modifications to the original instructional strategy that may be better suited to promoting student learning.

Teams:

- Reflect on common or disparate teaching experiences
- Identify alternative instructional strategies for future instruction
- Work with the content specialist(s) or leader(s) guiding the process to refine and improve the lesson
- Determine when the instructional modifications will take place. What can be built into subsequent lessons, and what needs an additional targeted lesson
- Record changes made to the lessons for future use

* *The Analyze and Adjust phases are dynamically linked and iterative.*

As Figure III.3 outlines, the PTLC had six phases. The first three, Study, Select, and Plan were usually addressed during the same meeting. Field staff would help teachers and other facilitators-in-training learn how to study the state standards. The criteria for selection of specific standards would usually be based on some existing analyses of student achievement data for the site. If the site had a written curriculum or scope and sequence, this document could guide the selection. If not, the group would select an appropriate assessment of the knowledge or skill expressed in the standard and prepare students for that. Once a standard or manageable set of them was identified, the next step was to determine how best to teach them. The group would talk about and decide upon which instructional strategies fit the lesson best. Field staff provided a lesson plan format to frame this discussion. Generally, by the end of the meeting, teachers would agree to implement this common lesson. In other words, they completed the fourth phase, Implementation, in between sessions. They also were supposed to bring back samples of student work from this lesson to the next session.

During the next meeting, participants would report back about their experiences with the lesson and think about ways to improve it. They might discuss how the instructional strategy could be used again in an upcoming lesson planned within the cycle or on their own. The cycle would be completed by engaging in activities relevant to the Analyze and Adjust phases. Participants would analyze student work, avoiding a focus on any individual student's performance. Rather, the goal of the analysis was to assess whether the lesson had been successful in teaching the standard to most students (i.e., whether they "got it"). Participants generated this information by looking at patterns across all of the student work samples to see what students grasped in general and what knowledge or skills they still lacked. During the Adjust phase of the cycle, participants strategized about how to integrate into future instruction these aspects of the standard the lesson had not successfully conveyed. The goal was to move forward with new material rather than teaching the entire lesson again. To design this new cycle, the phases would begin again at the next meeting.

The PTLC also built leadership capacity by promoting cross-level collaboration. Field staff encouraged principals and even district leaders to become involved in sessions with teachers or at least to visit their classrooms. As will be discussed here in the subsection on monitoring, the intention behind these visits was to collect data on the instructional program, not to evaluate individual teachers. Field staff built leadership capacity concurrently by instructing leaders to fulfill three functions related to the PTLC: (1) to set and communicate clear expectations to all staff about their participation, their roles and their responsibilities; (2) to monitor and review the activities during each phase of the PTLC, especially the actual Implementation phase; and (3) to provide support as needed to build the professional skills of teachers.

Competing School-Level Initiatives

Competing initiatives at the school level were a major obstacle to maintaining a coherent district course of action. This problem emerged when writing and enacting action plans. Frequent examples of these obstacles were school-level commercial programs, state-mandated school improvement plans, and state-level school improvement programs. For example, a site coordinator stated:

If a new program is introduced, the district leadership team can take a look at it to determine its worth. You'll remember that one part of Cube is creating coherence. That means making sure the programs are strong (. . .). Also, we need to monitor effectiveness in student achievement. If [a] program is not having the results expected, we need to determine whether it's [because of] implementation or the program itself. (Roydale, AR, 4/6/04)

As this field staff member alluded, field staff engaged in active investigation of the activities of other programs and professional development sessions offered to staff at participating schools to ensure that these services were high quality and aligned with the goals of the district action plan.

Table III.3 Other School-Level Initiatives Across the Sites

Site	Other School Initiatives
Athens, LA	LINCS (Learning Intensive Networking Communities for Success)
Bayou City, LA	LINCS (Learning Intensive Networking Communities for Success)
Bricktown, OK	Other TA provider ^a
Brookesville, OK	Literacy First
Desert Hills, NM	Reading First
Farmville, NM	Balanced Literacy
Forked River, AR	Reading First
Grisham, AR	None
Highway Junction, LA	LINCS (Learning Intensive Networking Communities for Success), Reading First
Pineland, TX	None
Roydale, AR	Reading First; Success for All; High Schools That Work; charter school
Wrightsville, OK	After school remedial program

Note. Since this list represents only what was reported in the SCRs examined for this report, a complete list of other initiatives active in participating schools during the course of the entire project would likely be more extensive.

a: This provider’s instructional strategies conflicted with the PTLC.

In some sites, field staff had to ensure that state-mandated school improvement plans would dovetail with district action plans. For example, during one of the Roydale, Arkansas district leadership team meetings, one of the agenda items was to go over connections between the district action plan and the schools’ improvement plans. The site coordinator helped the assistant superintendent review the district action plan. Then each building’s school improvement plan was distributed. The site coordinator asked representatives from each school to:

identify whether or not elements of the district action plan were evident in their school improvement plans. (...) School teams need to be attuned to evidence of implementation as well as evidence of the impact of the actions steps. While the district and school plans will not be identical, one should see a strong connection between the district and campus plans. (Roydale, AR, 12/6/04)

In other sites, school-level initiatives did not have goals incompatible with SEDL-affiliated work. Field staff still had to be aware of what was going on in schools throughout the

site and determine how the goals of systemic work could be mapped onto any independent but complementary initiatives. Such was the case for the Louisiana sites participating in LINCS (Learning Intensive Networking Communities for Success). Akin to SEDL's approach to school improvement, LINCS included regional coordinators who functioned like the field staff and "work[ed] directly with school and district staff, modeling lessons, observing and coaching classroom teachers, facilitating study groups, and otherwise providing individualized support to schools," (Louisiana Department of Education) The LINCS program also involved school improvement teams, faculty study groups, and content leaders.

According to a site contact record from one LINCS middle school in Bayou City, Louisiana, the topic of the October 17, 2003, campus meeting was to align LINCS with SEDL-affiliated work. The principal there had invited the SEDL site coordinator and site specialist to meet with the school faculty because the field staff had not met with them since interviewing them during the data collection done during the Site Entry/Data Scan stage. The site coordinator viewed this idea as a "transitional meeting" to get school staff on board with the SEDL-affiliated work and claimed it was a "great opportunity for reinforcing the idea that SEDL is not a program. This is a process and all levels are extremely important for a systemic school improvement." LINCS, in other words, was a resource that field staff could align with the SEDL action plan.

To make that point, the site coordinator reviewed the stages of the *Working Systemically* model, providing the middle school staff with the process the district leadership team used to develop the action plan. The site coordinator then pointed out that "the LINCS process has a school leadership team, and we would be working with the school leadership team and the math department to improve the curriculum, instruction, and assessment." As summarized in the site contact record describing these events, staff discussed a short article about other low-performing schools that had managed to be successful through "common strategies: take initiative, make no excuses, develop and execute a clear strategy, assess progress, research-based instructional

practices, and collaborate both inside and outside the school.” The purpose of presenting the article was to prepare the group for the next activity. The site specialist distributed information about specific steps needed to accomplish what was outlined in the action plan, and to start a discussion about how these activities would be monitored. To facilitate discussion, the site specialist divided participants into content areas and had them do a smaller scale needs assessment for their content area.

In a couple of cases, SEDL, in consultation with district officials, decided to withdraw from a school if it became clear that the leadership there had incompatible directions for its staff. For example, in Roydale, Arkansas, a charter school developed within a participating school that was already implementing Reading First and Success for All. Field staff decided that the district's decision to have this school pursue so many initiatives at once was incompatible with systemic work. In Bricktown, Oklahoma, field staff ceased attending school-level activities at the high school because the principal there did not want the faculty involved in the PTLC. Instead, that school was focused on the professional development offered by another consultant. Instead, field staff focused attention on Bricktown's middle school and district-level engagements. The site specialist for Bricktown reflected on the tension between independent initiatives and what SEDL was trying to accomplish:

This is the danger of using external TA [technical assistance] providers—we lose control over the content and focus of professional development. It takes diplomacy and additional effort to pull the pieces together and maintain the fidelity of the PTLC and the *Working Systemically* model (Bricktown, OK, 8/3/04).

In summary, written action plans were a core strategy field staff relied on to define the processes and procedures on which they would focus in the fourth stage of the model. These action plans were not always fully completed or approved prior to the sites beginning to enact parts of them. Where there was an action plan guiding work already begun, the field staff pointed out areas where the PTLC was compatible with the drafted plan and focused conversation and monitoring efforts on those components. Where the district leadership team was in the process of

writing a plan and the field staff had essentially already begun to engage educators in some practices compatible to the PTLC, they simply wrote it into the plan. The most challenging situations occurred when field staff had to try to reconcile the action plans' areas of focus with another existing initiative. Some of these sites and schools struggled along; SEDL dropped others because they were also pursuing other initiatives field staff considered incompatible with the *Working Systemically* model.

Building Educators' Capacity While Enacting the Action Plan

The field staff encouraged district and school leaders, prior to and as they shepherded the process of enacting the activities the action plan set out, to consider what specific skills educators in their sites would need to carry out the associated tasks. A central field staff strategy to support these leadership functions was to hold regular meetings with site-based teams, such as district leadership teams. Other teams were composed of only administrators or those that dealt only with school-level issues. The following discussion describes first what teams existed across the sites and outlines their respective roles. As a second strategy used to build leaders' capacity, field staff offered informal, one-on-one coaching, especially to superintendents and newly appointed district-level staff and principals. The primary way of building capacity at the classroom level was to engage educators in activities related to the PTLC, which is the third strategy analyzed. The PTLC strengthened both instructional leadership capacity and teachers' classroom instruction.

Site-Based Teams

Field staff strategy for developing, enacting, and monitoring systemic reform efforts was to work through teams of formal and informal leaders. Across all the sites, the most important team was the district leadership team. During the course of developing the model, however, field staff formed additional teams in the sites as needed. In some cases, schools already had teams meeting, and field staff incorporated their initiatives into the SEDL work to ensure they would be

congruent with it. Therefore, site-based team members were effectively the primary audience for most field staff strategies to build leadership capacity. Table III.4 outlines the types of teams field staff worked with across the sites.¹⁶

Table III.4 Site-Based Teams That Promoted Leadership Capacity

Site	District Leadership Teams	Administrative Teams	School-Level Teams
Athens, LA	X		
Bayou City, LA	X	X	X
Bricktown, OK	X		X
Desert Hills, NM			X
Farmville, NM	X	X	
Forked River, AR	X	X	
Highway Junction, LA	X	X	
Wrightsville, OK	X		
Brookesville, OK	X	X	X
Grisham, AR	X		
Pineland, TX	X		
Roydale, AR	X	X	X

District leadership teams.

The primary team throughout all stages of the model was the district leadership team, all of which were formed at SEDL’s initiative except in Forked River, Arkansas. Four sites relied solely on the district leadership team to direct the process. The size of these teams ranged from 7 to 21 members, depending on site needs and preferences. All included district leaders and principals, and, except in Athens, Louisiana, teacher representatives. Nine had active participation from the superintendent. In Bricktown, Oklahoma, the superintendent was not involved, but several high-level district administrators were. Pineland, Texas had only one

¹⁶ This table and analysis of site-based teams was developed from two sources of data. One was review of site contact records. Another was data collected from field staff during a series of discussions researchers organized with field staff on the topic of site-based teams. Three sites (Highway Junction, Louisiana; Bricktown, Oklahoma; and Grisham, Arkansas) also formed district-level content teams. Field staff formed these groups to address a lack of content knowledge among educators at all levels in the sites. These teams and additional groups formed to work on the PTLC are not included in the chart above, which focuses on teams that emphasized broader leadership issues and capacity.

district-level official on its district leadership team, and in Desert Hills, New Mexico, the field staff engaged only the school-level educators with no district or administrative leadership teams.

The district leadership team was the first audience for field staff's theoretical explanations of the *Working Systemically* model and the PTLC. Field staff interacted most closely with them in the earlier stages of the model to gather and analyze data about their schools. The district leadership teams wrote their action plans in close collaboration with the site coordinators. Throughout the process, field staff tried to model the professional conversations and collaboration style they hoped to see take root in the schools. Half of the district leadership teams agreed on explicit norms that bounded conversation allowed during meetings. At issue was not only that discussion should be restricted to items on the agenda, but also that the tone of the conversation was important. The idea behind the norms was to avoid or rein in "gripe sessions" at meetings that would only derail momentum for the work, especially when these lay blame on individuals. For example, in Grisham, Arkansas, the site coordinator "asked the group about what their expectations were for engaging in professional conversations about mathematics and also asked about establishing group norms. The group agreed that it would be important to have norms for their conversations and decided upon 'respect each other' and 'value opinions'" (Grisham, AR, 12/2/03). Such norms would be reviewed prior to each meeting (Highway Junction, Louisiana; Bricktown, Oklahoma; Grisham, Arkansas; Brooksville, Oklahoma; Farmville, New Mexico; Bayou City, Louisiana; and Roydale, Arkansas).

Primarily district staff had authority to provide resources necessary to complete the action plans and to decide how staff time would be allocated to associated activities in the requisite time lines and calendars. Field staff sought to influence these decisions through collaborative work with the district leadership team. District leaders were especially engaged in allocating resources to the PTLC in such sites as Brooksville, Oklahoma, and Farmville, New Mexico, where several schools were introduced to the PTLC, or where teachers received a stipend for their time to work on it, as in Athens, Louisiana; Highway Junction, Louisiana; and Pineland, Texas. The most

direct way field staff could align capacity-building efforts with the goals of the action plan was to be involved in delivering the district's professional development. This was the case by the end of the project for all of the sites that scored highly in the interview ratings for professional development (Athens, Louisiana; Highway Junction, Louisiana; and Roydale, Arkansas) discussed in more detail in Section IV. In all sites, at least some if not all of the professional development centered around the PTLC. Where other approaches to professional development initiatives conflicted with the PTLC, such as the high school in Bricktown, Oklahoma, SEDL withdrew from work with that site or those schools within it since their goals were incompatible with the field staff's strategy for developing systemic work.

Administrative teams.

Field staff initiated all administrative teams save that of Highway Junction, Louisiana. Administrative teams served practical functions related to the work of the district leadership team, such as helping with planning for district leadership team meetings and giving insight regarding "big picture" issues (e.g., district educational policy, personnel, resources). In addition, the team context provided a forum to address leadership accountability for actions taken and not taken. This check might not have occurred in district leadership team meetings where non-administrators were often present.

School-level teams.

Official campus-based teams existed in five sites: Roydale, Arkansas; Bricktown, Oklahoma; Brookesville, Oklahoma; Bayou City, Louisiana; and Desert Hills, New Mexico. They directed the process at the campus level.¹⁷ A review of site contact records from campus leadership team meetings for these sites revealed that discussions addressed interpreting school achievement data, suggestions for student interventions, professional staffing, communication

¹⁷ The function of these campus-based teams was distinct from those established in most sites to work on the more narrow goals related to learning how to do PTLC activities; these other kinds of groups will be discussed in the context of reviewing data regarding school-level professional development.

with parents, textbook adoptions, and other broader schoolwide issues. Field staff facilitated these meetings, offering professional literature on related topics and encouraging the teams to make decisions in the context of action plan priorities. Interestingly, the SEDL protocols did not call for campus leadership teams, and judging from a comment made by the site coordinator in Athens, Louisiana, they were not always necessary to the work. This field staff member wrote, referring to an existing school improvement team that was not chosen as a vehicle for SEDL work, that the field staff have “not had sufficient discussions about the role and function of the campus leadership teams. It seems that we may feel that it was an obligation to form one, but I really don’t know what to do with this group” (Athens, LA, 4/5/04). Even in Brookesville, Oklahoma, the formal campus leadership team met only three times, with an agenda item on two occasions being “Develop a plan to strengthen the functioning of the campus leadership team.” The item was not discussed at its first iteration because after a lengthy discussion about the importance of instructional leadership, the team ran out of time (Brookesville, OK, 12/9/04). In contrast, the campus teams led the work in Desert Hills, New Mexico, where due to several changes in district leadership, a team never took hold at the district level.

Coaching¹⁸

In addition to formal meetings, field staff relied on informal and continual one-on-one or small-group coaching to enhance leadership among district staff and principals. One-on-one conversations allowed leaders to talk about sensitive topics such as personnel issues (Farmville, NM, 4/27/05; Bayou City, LA, 5/4/05). Between visits, they would call and e-mail each other. Field staff often described their role as one of offering leaders both “pressure and support” (Highway Junction, Louisiana; Farmville, New Mexico; Forked River, Arkansas; Wrightsville, Oklahoma; Brookesville, Oklahoma). Or they advocated a “critical friend” role for themselves and other facilitators (Desert Hills, New Mexico; Brookesville, Oklahoma; Bayou City,

¹⁸ Data from this project regarding coaching strategies are also presented in last year’s report (Southwest Educational Development Laboratory, 2004).

Louisiana; Bricktown, Oklahoma), describing a role where someone offered constructive critique but did not direct team efforts. When they did not directly facilitate the district leadership team meetings, they were behind the scenes, helping to write agendas, to schedule meetings, and to create comfortable working environments so as to enhance productivity.

Another strategy was to arrange ad hoc meetings with leaders. This form of communication was especially important when there was leadership turnover. During academic years 2003–2004 and 2004–2005, there were new superintendents in Bayou City, Louisiana, Farmville, New Mexico, Wrightsville, Oklahoma, and Desert Hills, New Mexico, with a rumored impending change in Pineland, Texas. The site coordinator in Bayou City, Louisiana, reported that field staff “met with the new superintendent to receive his support for continuing the SEDL work.” They drew up a “2003/2004 calendar for SEDL’s visits to the parish and creat[ed] the available time to interact with both the schools/district team.” Field staff also had to concentrate on building relationships with newly assigned principals so as to maintain understanding of and momentum for the plan “at the top.” Twelve of the 28 schools had new leadership during the last 2 academic years. Principals otherwise were not often direct targets for individualized coaching. Field staff interacted extensively with them during team meetings because principals would be present at district leadership, school-based, and some administrative team meetings.

Field staff invited district and building leaders to attend SEDL’s leadership institutes, held in the summers of 2003, 2004, and 2005. Leaders gathered for 2–3 days in SEDL’s Austin office for formal training in such arenas as instructional leadership, the practices of high-performing school districts, leaders’ roles in implementing and monitoring the PTLC, and strategies for sustainability. Field staff made explicit reference to following up on leadership institute training in site contact records from Wrightsville, Oklahoma, Desert Hills, New Mexico, Athens, Louisiana, Forked River, Arkansas, Roydale, Arkansas, and Bricktown, Oklahoma. For

example, a requirement for attendance during summer 2005 was a completed action plan for academic year 2005–2006.

In summary, field staff used two primary strategies to promote instructional leadership. One was to guide the work by modeling leadership while facilitating site-based team meetings. The other was to offer ongoing coaching to official leaders, especially superintendents whenever possible. Obviously, their commitment to the work was crucial to making sure resources were available and motivating other leaders in the district. An important goal of strengthening instructional leadership was to bolster administrators' abilities to communicate clear expectations to teachers regarding the PTLC. Principals were crucial to monitoring teachers' professional development, so they were also included in field staff coaching efforts, though less often in the one-on-one meetings characteristic of field staff approaches to district leaders.

Approaches to the PTLC

In reviewing the site contact records and all responses to the question “What did your district or school do to implement the PTLC?” in the spring interview, there appeared to have been three different approaches to the PTLC, which sometimes overlapped:¹⁹

- SEDL provided large-scale professional development that supported the PTLC, which was facilitated between field staff visits primarily by site leaders or teachers.
- SEDL facilitated small-group PTLC sessions while on-site and also provided professional development to literacy or math coaches (and/or district-level content experts), who were primary facilitators of the PTLC when SEDL was not on-site.
- SEDL facilitated small-group PTLC sessions with some support from site leaders in between visits. Most of these site leaders did not have content expertise.

The site specialists delivered large-scale professional development, most often at the request of a district-level curriculum staff member who asked them to lead daylong professional development sessions, such as those that commonly kick off school years. Field staff used this

¹⁹ This discussion of the PTLC identifies general trends research identified when reviewing site contact record and interview data, which were both collected at the site, not school, level.

time to give contextual background to the longer processes in which sites were involved (such as how the PTLC fit into the “big picture” of systemic work or to model an entire cycle of the PTLC). SEDL delivered this kind of professional development to entire district faculties in Roydale AR, Highway Junction, Louisiana, Farmville, New Mexico, and Brookesville, Oklahoma.

Within these sessions, field staff frequently broke the audience into smaller groups to promote collaboration. For example, in Highway Junction, Louisiana, the educators devoted one half-day a month to the PTLC. Field staff led an introductory general session for all teachers concerning the PTLC or reviewed particular phases of it. Into this presentation, they incorporated some specific instructional strategies related to a strand of the state standards that addressed the component prioritized by that site's action plan. Then small groups would work on collaborative lesson development, share their lessons, and brainstorm about ways to analyze them, with a wrap-up focused on deciding on next steps. As the day closed, the teachers would make plans about when they would teach and assess the lesson and when they would regroup and analyze their students' work later (for a detailed example of this process in another site, see record for Brookesville, OK, 12/9/04). The site specialist in Pineland, Texas, rather emphatically asserted after working in this manner with all teachers for grades 1–5: “I would like to revise the PTLC ‘how-to’ manual with a bold-face statement declaring that **THIS PROCESS DOES NOT WORK WELL WITH LARGE HETEROGENEOUS GROUPS**. It was made worse by the absence of the principals” (Pineland, TX, 8/12/04).

Small group work was the focus of two other approaches to the PTLC. The chief difference between the approaches was whether the on-site facilitator who kept the process going in between visits from field staff had content expertise. When field staff did not directly offer content expertise, they trained professional staff with content expertise, such as curriculum coordinators or math and literacy coaches, to facilitate the PTLC sessions in their absence. Examples of this approach were the site work in Desert Hills, New Mexico, Athens, Louisiana,

and Grisham, Arkansas. The field staff in Grisham, for instance, met with math coaches and the district director of curriculum and instruction to help them with content and pedagogical knowledge as well as facilitation skills. The coaches reported on what they had been doing with the teachers, and field staff assisted in identifying which strand of the state standards would address student deficiencies as identified by the coaches' analysis of their district's state-mandated test data (Grisham, AR, 9/1/04). The site coordinator believed that the coaching structure that developed in Grisham had "the potential to provide ongoing, job-embedded professional development to support teacher growth"—language that was very compatible with the *Working Systemically* model's competency of "ensuring continuous professional learning." According to this field staff member, the coaches provided a "critical" link across levels of the system because "they participate in DLT meetings, school meetings, and work in the classroom with teachers" (Grisham, AR, 9/17/04).

In the third approach, on-site facilitation was left in charge of principals and teacher-leaders, who were generally not content experts. Examples of this delivery method were found in records for Roydale, Arkansas, Pineland, Texas, Bayou City, Louisiana, Bricktown, Oklahoma, Brookesville, Oklahoma and Farmville, New Mexico. The site contact records are thin on reports of how the work actually proceeded in these groups when the field staff were not on-site. Debriefing with principals was folded into meetings where several other issues might be on the agenda. In peer debriefing about this part of the analysis, one field staff member expressed concern about how the work was going. These opinions were informed by conversations with educators on site.

I am pleased that [educator] recognized her teachers were not truly collaborating. Most administrators in the sites tell us that their teachers are "doing it" and accept the evidence of graphic organizers as proof of the PTLC. [Educator], however, approached us about doing more than just completing organizers. She was concerned about true collaboration and was begging for our involvement to improve the process and ensure fidelity to the PTLC. This was rare. (Roydale, AR, 1/18/05)

Challenges Associated With the PTLC

All three of these approaches ran into some problems. No matter the mode of delivery, the district level could help or hinder progress depending on what kind of curricular and assessment resources it made available. As the site specialist in Farmville, New Mexico, pointed out, the PTLC depended on many other resources, such as “a scope and sequence,” “reliable measures of student progress,” “time to do the work,” appropriate assessments, and interventions. The tasks outlined in the action plans generally addressed these needs. The field staff helped district leaders direct processes to develop a district curriculum, scope and sequence, and benchmark assessments aligned to the standards. These were usually still works in progress.

Another district-level issue was deciding when and how often the collaborative time necessary for the PTLC would take place. One advantage to the choice of delivering the PTLC on a large scale during teachers' formal professional development time was that it eliminated complaints teachers made about having to give up their common planning time for PTLC work. Most resistant were teachers who were pulled out of class or had to stay after school, as was the case in some schools in Pineland, Texas. The field staff response was to work with district leaders to ensure that grade-level or subject-area teachers had common planning time together during the day. While on-site, field staff would attend these meetings and point out how developing a portfolio of common lessons lessened the responsibilities of individual teachers. Some sites also decided to compensate teachers for any time they were asked to spend after school (Forked River, Arkansas; Athens, Louisiana; Highway Junction, Louisiana; Pineland, Texas). In Farmville, New Mexico (elementary), Bricktown, Oklahoma (middle school), and Pineland, Texas (primary school), field staff reported that faculty of at least one school were regularly participating in PTLC activities in engaged ways. In all of these schools, teachers collaborated at least once a week.

Interestingly, none were high schools. In interviews with high school teachers, some reported that they found it difficult to use their common grade-level planning time fruitfully

because they perceived their subject matter as too diverse to benefit from the collaborative effort. SEDL worked with high schools in Farmville, New Mexico, Brookesville, Oklahoma, Bayou City, Louisiana, Forked River, Arkansas, and Roydale, Arkansas, and in four out of five of these schools, field staff and educators alike reported that the PTLC had not taken root. Common planning time was scheduled, but teachers did not use it effectively or at all for the PTLC. In short, field staff strongly encouraged districts to schedule common planning time for the PTLC groups to meet; having this time was necessary but not sufficient for successful work, especially at the high school level.

There were advantages and disadvantages to each of the approaches with respect to the PTLC itself and to the capacity levels of those who were participating in it at the school level. The primary advantage to the large-scale professional development approach was that field staff controlled the information provided to teachers. They were also able to circulate among the small groups and get an idea of how well educators understood what they were doing. In all three sites where large numbers of faculty were being trained on the PTLC, however, this opportunity to see firsthand what the small groups were doing led to concerns about how well they grasped the connections between the phases of the process.

In the second approach, where math or literacy coaches facilitated the PTLC in small groups, lack of facility with the standards, content knowledge, and pedagogical skills were less of a problem. The cycle nonetheless was difficult even for coaches in these sites to grasp. As one field staff member talking about Desert Hills, New Mexico, put it, “It is frustrating that even eager and capable coaches still struggle” with the PTLC. For example, in Highway Junction, Louisiana, the district had adopted the *Voyager* reading program. Teachers used that program as their starting point when designing lessons, rather than becoming familiar with the standards via studying them directly and then looking over available instructional resources to fit those to their lessons. Across the sites, teachers had textbooks, favorite lessons, and activities that they also

used as starting points, hunting through standards to match the instruction with which they were already comfortable or which was already structured for them.

To encourage educators to start with the standards when planning lessons, field staff developed a lesson plan format that required teachers to specify how the lesson connected to the standards. They asked educators to explain how their selection of a lesson's standards connected to student needs. Field staff in all sites modeled lessons tied to state standards for teachers so that they could see what such a lesson would look like. During the Analyze phase of the model, even in sites where there were content coaches and a scope and sequence, such as Grisham, Arkansas, educators did not always understand that the purpose of analyzing students' work was to help make decisions about future instruction. A common strategy during the Analyze stage to help session participants focus on the lesson's overall impact (rather than on the performance of individual students) was the "three-stack" method, wherein names of students would be concealed and their work sorted into three piles: below proficient, proficient, and above proficient. Teachers could then examine each stack to identify where the lessons had been successful or not successful. Conversation would then be about what students across the instructional program still needed to know, leading into the Intervene phase.

Since even the sites that utilized content experts as facilitators experienced problems related to capacity, there was not a discernable difference in the reports of low capacity for sites. Rather, the driving factor in the sites that relied primarily on principals and teacher-leaders to facilitate the PTLC sessions between SEDL visits was whether the principal became involved. The advantage to leaving principals in charge was that they were in a position of greater authority and could elicit more accountability among teachers charged to implement lessons. The disadvantage was that if the principal did not actively monitor or build capacity among teacher-leaders, teachers often did not know what to do or chose not to collaborate at all. In the site contact records for 10 of the 12 sites, field staff credited or blamed the role of the principal for the success or failure of the PTLC. For example, in Farmville, New Mexico,

Where the principal was involved, facilitating and setting expectations, the PTLC took hold. It's not perfect there yet, but it is underway. Where the principals looked for other facilitators or allowed managerial items to enter into the time, the PTLC did not take hold. The leadership roles appear to be the key. (Farmville, NM, 4/27/05)

Despite these problems, those educators who did become engaged in the PTLC were positive about its value overall. For instance, teachers said that they liked having a process to follow and enjoyed the opportunity to get new instructional ideas from colleagues, although some remained concerned about being able to retain their own individual teaching style. For some, the notion that grading had a purpose beyond individual student evaluation was an eye opener. Given that their students had to take the state-mandated tests, they wanted their lessons to be tied to the standards to help their students succeed. They also appreciated that the PTLC centered attention on student achievement.

Their main concern was that the PTLC was very time-consuming in two ways. One was that it took going through it for a few cycles until educators fully grasped the connection among all the steps. The other was more literal; teachers wanted more hours per week devoted to the collaborative process. In the second and third ways of approaching the PTLC, it was common during the project for only a few groups to be asked to pilot it. These schools uniformly reported that they wished more faculty had been able to be involved. A third concern was that teachers in schools where principals did not get involved felt less accountable to the work, which was frustrating for those who were more engaged.

To sum up, educators were positive about the idea behind the PTLC but had concerns about its feasibility, especially in the absence of support and monitoring. Field staff did not consider the PTLC inherently problematic but recognized it was not a set of practices educators could or would immediately adopt. Because scheduling time for the collaboration was a barrier in some sites, field staff tried to cooperate with those sites that preferred the PTLC work to be scheduled during in-service days. Whenever possible, however, the field staff facilitated small collaborative groups where they could build participants' content knowledge and pedagogical

skills more intensively. This choice left more of the monitoring to principals, who were not always up to the challenge of functioning as instructional leaders when the field staff were not on-site.

Monitoring

Field staff encouraged monitoring of two domains of activity. One was of the PTLC activities. To promote this, field staff in all sites encouraged leaders to attend teachers' work sessions. All field staff also asked leaders to conduct the classroom observations that have come to be known as "classroom walkthroughs," or "CWTs." A third, less formal monitoring practice was "one-legged interviews." Descriptions of these practices will be examined in turn in this section.

The other area requiring monitoring was the overall progress and timeliness of action plans' progress toward action plan goals. Tools that field staff introduced to monitor overall progress of the action plans will be described: the Innovation Configurations (ICs) and Action Plan Implementation Reviews (APIRs). Where these tools were not adopted, field staff reviewed progress by including the action plans as a standing meeting agenda item. Finally, what was learned via these monitoring efforts will be summarized.

Monitoring PTLC Activities

Attending teachers' collaborative work sessions.

The field staff asked leaders, especially principals, to be actively engaged in instructional collaboration with teachers. One field staff member described such efforts:

I clarified the administrator's role in the PTLC process: "The principal's role is to listen if that lesson that is being planned is really aligned to the standards. Principals should be asking clarifying questions. They are not participating as an up front facilitator, but they're to ask probing questions, trying to build teacher capacity. The administrators are simply an ear, listening to what is happening." (Brookesville, OK, 12/9/04).

Helping leaders find the time to attend their faculty's collaborative meetings was an uphill battle, yet most leaders recognized the value of monitoring. The comments of one elementary principal in Pineland, Texas, illustrated this belief: "In hindsight, I should have attended those meeting[s], but I had to get our campus plan done. I now realize that I need to talk more with the teachers about this process and what we are doing" (Pineland, TX, 10/28/04). As the comment suggests, however, principals did not always realize the importance of their physical presence until it was too late. Throughout the project, field staff struggled with getting leaders to prioritize instructional leadership over management issues, and asking them to attend the PTLC session was part of this difficulty.

Classroom walkthroughs.

Brief classroom observations via classroom walkthroughs, with subsequent constructive dialogue between leaders and teachers, were a primary vehicle for encouraging cross-level cooperation as well as sustainability for the work. These brief classroom observations were promoted in all sites as part of the leaders' role in monitoring PTLC progress. Field staff hoped district and school officials might begin to view themselves as instructional leaders, not just administrators. The classroom walkthroughs helped leaders gauge the quality of their schools' instructional programs.

All field staff and one researcher were trained how to do classroom walkthroughs themselves in December 2004. Some sites had received training from other sources, such as 24/7 in AR. Where that was not the case, the field staff often accompanied superintendents, principals, and other leaders on walkthroughs and pointed out what to observe. They also developed checklists to monitor progress or modified existing versions of such lists already in use. The site contact records for Bayou City, Louisiana, provide the following example of how field staff approached training leaders on the classroom walkthroughs:

The administrators were asked to visit two math classrooms at [High school] and write down everything they could observe with a short four minute [visit]. The

administrators returned to the conference room with their observations. The administrators were asked to work in their groups and to categorize their observations and then label with a topic. Then as a large group [they] agreed upon the categories and place stickers in the appropriate area. (. . .) Examples of the categories were as follows: 1. Use of technology (Resources) 2. Teacher engagement 3. Board Assignments 4. Visuals Present 5. Assistance and 6. Student Engagement. (Bayou City, LA, 1/11/05).

As happened during the field staff classroom walkthrough training, observation tended to be more on the environment and the students than upon factors that related to student achievement. District officials and especially principals needed to know what criteria to look for as signs of quality instruction, such as students being aware of what a lesson's objective was, having that objective be explicitly tied to a grade-level standard and delivering that objective through materials and instructional strategies that would engage students in higher-order thinking.²⁰ The site coordinator in Bayou City was pleased with the tenor of the discussion held in the aftermath of the classroom visits because instead of talking about "checking lesson plans, documentation for LINCS, discipline etc.," she felt that the leaders were recognizing that "effective instruction limits off-task behaviors, lack of motivation, etc. (. . .) There was no talk about discipline or blame." (Bayou City, LA, 1/11/05)

One concern leaders had with walkthroughs was that their presence would make teachers anxious. Field staff prompted leaders to let teachers know that the walkthroughs were going to happen and to make it clear that their purpose was to collect data about progress toward alignment goals, not to evaluate individuals' instructional performance. As a site coordinator in Highway Junction, Louisiana pointed out:

The intent was not to evaluate teachers on these walkthroughs but to get a sense of instruction across the school to assist with decision making and to practice providing teachers with questions geared to having them reflect on their practice. (Highway Junction, LA, 2/16/05)

²⁰ Many field staff referenced Robert Marzano's work on "high-yield" instructional strategies (2001).

The field staff also coached leaders on how to follow up their visits with constructive informal chats with teachers.

[Principal] was primarily concerned about her teachers' receptiveness. She wanted to address the issues we observed in the CWTs but not single out any teachers or put them on the defensive. She wanted them to get them [sic] studying the standards and connecting the content to the comprehension standards, but she didn't want to ask them to do too much work. She wanted to expand the PTLC teams by using the current team members as facilitators, but she didn't want any of the teachers getting defensive about having to do "another thing." (Bricktown, OK, 1/12/05)

As this site contact record describes, the site specialist coached her through planning an agenda for the faculty meeting. She "would open with a reminder of the purpose of CWTs. She would then discuss the pattern observed in the data from the first CWTs: a lack of understanding about what it is students should know and be able to do." The principal then would assign departments a standard to study from a list pre-selected in collaboration with the SEDL site specialist. The selection criteria were those that "were particularly open to interpretation so as to highlight the need for collaboration" and, based on data analyzed from prior visits, on a standard "that we observed being taught at a lower level in seventh grade than in sixth." Departments would then "compare their analysis of the standard to what they wrote on the curriculum maps" at a professional development session held prior to the opening of the school year. The principal would close the meeting with a reflective question: "Where the standards are vague or repeat from grade to grade, how do you determine if your lesson is on target?"

Judging from a review of interviews, teachers generally were not apprehensive about leaders visiting their classrooms. In contrast, except for the teachers in Brookesville, Oklahoma, who did object, teachers appreciated the interest taken in their instruction and felt leaders could make better decisions when they were more aware of what was going on in the classrooms. A bigger challenge to the process was for those who conducted walkthroughs to learn how (and to find the time) to analyze the data collected during the visits. Researchers found examples of data being analyzed by educators only at field staff prompting.

One-legged interviews.

One-legged interviews were a third strategy for monitoring PTLC activities. Hall and Hord (2001) describe the process as follows:

One of the major types of small interventions is what we call the “one-legged interview.” One frequent opportunity for one-legged interviews occurs when a teacher and a principal meet in the corridor. If the principal or teacher initiates a brief discussion about the innovation, then it is a one-legged interview type of intervention. We use the name “one-legged” to indicate that these interventions are brief (most people can’t stand on one leg very long.), since both the teacher and the principal probably have to be somewhere else when the next bell rings. Yet a moment was taken to talk about the teacher’s involvement with the innovation. (p. 10)

Such “one-legged interviews” were part of field staff strategy in Forked River, Arkansas, Brookesville, Oklahoma, and Highway Junction, Louisiana, and advocated for use by superintendents as well as principals. Concerns raised about this technique were similar to those related to classroom walkthroughs: leaders were unclear what exactly to address, had trouble finding time to do them, and did not often think systematically about what they were learning when they did have these casual conversations with staff.

Monitoring Progress Toward Action Plan Goals²¹

Two prominent tools field staff employed to monitor progress toward goals set out in the action plans were Action Plan Implementation Reviews (APIRs) and Innovation Configurations (ICs). An Action Plan Implementation Review was a simple chart on which to record progress toward the completion of the action plan. During the Planning Action stage, only the Description of the Ideal State, Critical Element, and Action Steps were included in the action plan. Educators completed the rest of the document during discussions at district leadership team meetings. See Figure III.4 for a sample Action Plan Implementation Review form that was a part of the field staff protocols for this stage of the model.

²¹ This section describing the APIR and IC is adapted from a 2003 version of the project protocols.

Figure III.4 Sample Action Plan Implementation Review

Date of Review: _____ Date of Next Review: _____
 Description of the ideal state: Everyone understands the district curriculum that is aligned with state standards and uses it to plan instruction
 Critical Element: Development of a district reading curriculum that is aligned with state standards

Action Step	Evidence of Implementation	Evidence of Impact	Still need to do
Create a reading curriculum team composed primarily of teacher representatives from all levels.			
Central Office makes and distributes copies of state reading standards to all teachers and principals in district.			
Provide training to curriculum team on curriculum alignment.			
Curriculum team compares specific objectives from state standards for each grade level with curriculum being taught in district schools.			
Align reading curriculum to objectives/benchmarks from state standards for each grade level across all grades K–12.			

Members of the team provided updates on their activities since the team had last met and discussed what concerns they had with respect to their progress. They then revisited and adjusted the time line. The field staff made sure to link activities back to the importance of alignment and improving students’ performance. Additional work might have been added to the action plan. At the end of the meeting, the field staff encouraged members to reflect on what they were learning during the process of enacting the plan.

The field staff also found it helpful to have a document that painted a clear picture of specific behaviors that each person must do to accomplish the plan (What do district staff members do? What do principals do? What do teachers do?). Some used an Innovation Configuration (IC) to clarify behavioral expectations and monitor implementation of the action plans (see Figure III.5 for an example taken from the project protocols).

Figure III.5 Sample Innovation Configuration

Ideal				Unacceptable
Component A: The superintendent makes the curriculum alignment and use a district priority.				
1	2	3	4	5
The superintendent makes the curriculum alignment work a high priority in the district as evidenced by providing adequate time and funding and making it a regular topic at board meetings, principal meetings and other administrative meetings and all district-wide events.	The superintendent makes the curriculum alignment work a priority in the district as evidenced by occasionally referring to it at board meetings, principal meetings and other administrative meetings and all district-wide events. Some time and money are allocated to the work.	The superintendent makes the curriculum alignment work a priority in the district as evidenced by providing adequate time and funding. There is no mention of the curriculum work in meetings or significant events.	After the initial introduction of the curriculum, reference is made to it at district meetings or principal meetings at the beginning of the year but not again. Limited time and funds are set aside to support the work.	After the initial introduction of the curriculum, no reference is made to it at district meetings or principal meetings. No time or funds are set aside to support the work.
Component B: Central office staff identifies the specific knowledge and skills that students will acquire by grade and subject.				
1	2	3	4	5
Central office provides all teachers with a document that supports the state standards by identifying the specific knowledge and skills that students will acquire by grade and subject. Teachers have the documents for grades that come before and after theirs.	Central office provides all teachers with a document that supports the state standards by identifying the specific knowledge and skills that students will acquire by grade and subject. Teachers have the documents for their grade only.	Central office provides all teachers with a document that supports the state standards by identifying the general knowledge and skills that students will acquire by grade and subject. Teachers have the documents for their grade only.	Central office provides all teachers with a copy of the standards with directions that they are to be referenced in lesson plans.	Central office provides all teachers with a copy of the standards.

Field staff supported educators as they tailored the IC to their action plan. Together, they determined the major components within the action plan (e.g., implementation of professional development, what teachers were expected to do in their classrooms as a result of the professional development, or how new structures to support teachers were being carried out). Field staff recommended selecting only five or six major components because too many make the process too burdensome. The team would then develop rubrics related to these key

components and associated activity, describing 3–5 variations of implementation ranging from perfect to unacceptable. A draft of the IC was shared with site leaders and they had the opportunity to revise it.

Table III.5 shows that not all the sites used these tools, nor were they uniformly developed across the sites during this stage of the model. More sites may have used these tools during work prior to this time period, or field staff may have used the tool but not reported its use in their site contact record.

Table III.5 Tools for Monitoring Implementation of Action Plans

Site	Action Plan Implementation Review	Innovation Configuration
Athens, LA		X
Bayou City, LA	X	X
Bricktown, OK		X
Brookesville, OK		
Desert Hills, NM		
Farmville, NM		
Forked River, AR		X
Grisham, AR	X	
Highway Junction, LA	X	X
Pineland, TX		
Roydale, AR		
Wrightsville, OK		

Field staff for three sites noted various issues related to difficulties using this tool. In two of these sites, educators were uncomfortable evaluating the process, especially on issues related to leadership because that would reflect directly on the performance of the superintendent himself. In one of these two sites, as well as in a third site, the field staff did not find them helpful because the educators spent too much time trying to figure out the components and descriptors and were not able to implement the tool for evaluative purposes.²²

²² These three sites are not named to ensure the anonymity of the superintendents.

Informal review of progress on action plans.

In all sites, review of action plans was a standing agenda item at district leadership team meetings. With or without use of the formal tools, field staff engaged in the more informal strategy of ongoing dialogue about progress. Field staff held team discussions about what was in the plan, making sure participants raised concerns in order to check levels of buy-in. The discussion also attended to fleshing out time lines, noting associated tasks and parties who would carry these talks out. Field staff helped leaders in all sites analyze student achievement data and identify areas that should be priorities and presented this information to teachers to solicit their support.

Issues Identified via Monitoring Processes

The site contact record template did not include a field for the APIR or IC, so data regarding their findings were not systematically recorded. Review of the site contact record data where field staff happened to bring them up suggested the kinds of issues these tools could identify for those sites that used them. Sites that assessed their own monitoring efforts, such as Bayou City, Louisiana, and Highway Junction, Louisiana, knew that they were not performing this function very well.

In Brookesville, Oklahoma, review of the action plan led the district leadership team to conclude that they needed to spend more time learning how to implement new instructional strategies in their classrooms. In Athens, Louisiana, the district leadership team identified a need to develop benchmark assessments, with the recommendation that they be piloted first with accompanying professional development for staff regarding the use and benefit of the assessment; how to administer it; how to analyze the data it would generate; and how to hold team or grade-level meetings to debrief and plan interventions for students based on the data. The group decided that the actual benchmark assessment should replace an existing test. They thought some test items should be formatted in a way similar to standardized tests used in the

district and that the revised exam should cover all appropriate parts of the state standards for a given subject or grade. In Highway Junction, Louisiana, review of the IC results led to a similar discussion regarding their need for new assessments to monitor student progress. Field staff helped team members disaggregate their student achievement data and identify relevant trends that might suggest new courses of action.

Because most sites did limited amounts of monitoring, they were not able to make adjustments to action plans based on data they had collected themselves. However, the monitoring tools and techniques did have positive effects on decision making processes. Because some leaders in all sites did classroom observations at least when in the company of field staff, they became somewhat more familiar with their instructional program's quality and rigor even if they did not usually collect systematic data using this technique. Similarly, the IC and APIR tools had some utility even though they were generally not used as intended because trying to develop one stimulated discussion about what the team members thought were necessary adaptations to the action plans. Most typically, informal reviews of activity that occurred between field staff visits informed updates to action plans, adjustments to the action plan's time lines, and deletions of tasks as they were accomplished. To help make decisions about these revisions, field staff also disaggregated district student achievement data for the SEDL teams and distributed relevant professional literature from current educational research.

The strategies for monitoring classroom-level practice were designed to build relationships across levels of the educational system, particularly between principals and teachers. They required cross-level communication about key components and competencies. Monitoring progress toward the goals set out in the action plans was another way to ensure alignment throughout levels of the system. Most sites reached the point where leaders recognized the value of these monitoring techniques and tools but were still not comfortable enacting them or reported that they did not have time to do them.

Discussion

As is evident, developing systemic work is complex and necessarily involves multiple, simultaneous, and ongoing strategies. A tension that ran throughout the entire project was how to balance progress toward the goal of increasing student achievement, which required tailoring activity to site needs, with maintaining consistency across the sites in how to work systemically. Certain core strategies were nonetheless apparent. First, the field staff encouraged leaders to identify clear goals by writing them down in action plans. Second, as these goals were implemented, field staff spent time building educators' capacity. Systemic work involved the collaborative work of cross-level teams. During these settings and in less formal ways, field staff coached leaders on how to perform the role of instruction leader. Alignment of curriculum, instruction, and assessment to the state standards was theorized to have the most direct relationship to improving student outcomes. The PTLC was thought to be the best way to create that form of coherence as well as to build capacity at the classroom level. Third, careful monitoring was required to keep track of PTLC activities and of overall progress districtwide toward other goals expressed in the action plans. Leaders were encouraged to monitor PTLC activities by attending teachers' sessions, visiting classrooms, and communicating more often and less formally with teachers. Another important area of monitoring was to review progress toward districtwide goals using whatever implementation and student achievement data were available.

All three core strategies targeted strengthening different competencies related to systemic reform. To do so may require more intensive and longer-lasting support, especially for facilitating the PTLC, as field staff suggested above. One site coordinator was especially frank, stating the following:

The PTLC is not easily implemented. It seems to take numerous demonstrations and extensive professional development. For schools and districts that have no prior experience, the alignment process is arduous. We often assume that there is capacity to carry out steps that doesn't actually exist. The optimal situation would be for SEDL staff to be on site constantly during the first phases of the alignment

work to answer questions, to iron out details, and to offer immediate support (Brookesville, OK, 8/23/04).

Judging from field staff reflections and some of the interview comments, such a promising but complex initiative might have best been implemented by an on-site facilitator. Because none of the sites received such intense facilitation from field staff, both field staff and educators looked to on-site leaders for direction. Fullan's (2001) comment regarding the importance of the leadership at "one level up" of a particular part of the intervention, cited earlier in this report's literature review, also seems pertinent. Because the project engaged proportionately more of the district leaders, the superintendent's involvement was especially appreciated (or missed where it was absent), according to both the site contact records and interviews. Similarly, given that the PTLC was a classroom initiative, principals had a crucial role to play.

Without strong leadership to continue directing the systemic work, field staff had doubts that their efforts would be sustainable. Sustainability is not a dichotomous variable, however, and educators reported intentions to continue many central aspects of the model. Whether they can do so without the support of the field staff remains to be seen, as does whether those sites that do continue the work will eventually show greater gains in working systemically and student achievement.

IV. WORKING SYSTEMICALLY OUTCOMES

To what extent did districts and schools increase their capacity to work systemically?

To address this research question, the following two analytical subquestions were posed:

- Did sites increase their systemic work over time?
- Were there differences in development of systemic work?

Fall interview and *Working Systemically* survey data were used to determine whether sites increased systemic work over time. Qualitative analyses of selected sites were also conducted to elaborate on sites with notable results and to develop possible explanations for differences in results among those sites. Results for the two subquestions are presented first, followed by results of qualitative analyses of selected sites. The section concludes with a discussion of findings. Section II describes in detail the samples, data sources, and measures used.

Findings

Increases in Systemic Work Over Time

To determine whether systemic work increased over time, descriptive data for both interview and survey scores were examined. Statistical analyses (*t* tests) were also conducted to determine the significance of changes in cross-site mean scores on both measures over time (comparisons of sites across time are reported in the following section). An alpha level of .05 was used in all statistical analyses.

Interview Results

As described in Section II, interviews were assigned ratings on four rubrics measuring progress in sites' work on expectations and priorities, alignment, collaboration, and professional development. These ratings were summed to create overall interview scores for each site in 2003

and 2004. The following table reports the mean site sum scores, standard deviations, and change scores for each site. Possible scores ranged from 4 to 20.²³

Table IV.1 Descriptive Statistics for Interview Scores, 2003–2004

Site	2003			2004			Change
	N	Mean	Std. Deviation	N	Mean	Std. Deviation	
Athens, LA	8	9.50	1.512	8	14.13	2.232	4.63
Bayou City, LA	7	12.14	2.268	7	13.00	3.055	0.86
Bricktown, OK	9	8.56	1.130	9	11.56	2.297	3.00
Brookesville, OK (T)	7	8.57	2.637	8	9.88	1.126	1.31
Desert Hills, NM	8	11.13	3.137	8	12.00	1.309	0.87
Farmville, NM	7	9.43	1.512	7	11.29	2.430	1.86
Forked River, AR	6	9.50	1.643	8	11.13	2.100	1.63
Grisham, AR (T)	6	8.50	2.258	6	12.00	1.414	3.50
Highway Junction, LA	8	10.63	1.923	9	12.89	.928	2.26
Pineland, TX (T)	8	11.88	1.642	8	12.88	2.949	1.00
Roydale, AR (T)	8	9.13	2.780	6	13.17	2.563	4.04
Wrightsville, OK	5	11.60	1.140	7	9.86	1.773	-1.74
Total	87	10.02	2.338	91	11.98	2.343	1.94 *

T: Test site

* $p \leq .001$

Table IV.1 shows that scores in all sites but one increased from 2003 to 2004. Across sites, scores rose by 1.94 points. Increases were also found in scores for each of the four rubrics. Tables IV.2–IV.5 report the site mean scores, standard deviations, and change scores for all sites for each rubric. Mean scores were used as opposed to sum scores to account for differences in sample sizes between sites. Possible scores ranged from 1 to 5. As expected, sites showed more variation in their scores on individual rubrics than on mean interview sum scores.

²³ A score of 4 represents a rating of 1 on each of the four rubrics. A score of 20 represents a rating of 5 on each of the four rubrics.

Table IV.2 Descriptive Statistics for Expectations and Priorities Rubric

Site	2003		2004		Change
	Mean	Std. Deviation	Mean	Std. Deviation	
Athens, LA	2.38	0.744	3.75	0.463	1.38
Bayou City, LA	3.00	1.07	3.86	0.900	0.86
Bricktown, OK	2.89	0.601	3.56	0.726	0.67
Brookesville, OK (T)	2.71	1.11	3.38	0.916	0.66
Desert Hills, NM	3.13	0.991	3.38	0.518	0.25
Farmville, NM	3.14	0.378	2.71	1.254	-0.43
Forked River, AR	2.83	0.753	3.38	0.518	0.54
Grisham, AR (T)	2.83	0.753	3.67	0.816	0.83
Highway Junction, LA	3.25	0.707	3.44	0.882	0.19
Pineland, TX (T)	3.50	1.25	3.50	0.753	0.00
Roydale, AR (T)	2.88	0.926	3.83	1.069	0.96
Wrightsville, OK	4.20	0.447	2.86	0.690	-1.34
Total	3.03	0.895	3.44	0.833	.41 *

T: Test site

* $p \leq .01$

The highest final year (2004) mean was found for the Expectations and Priorities rubric. This makes sense given that the first steps in working systemically were to identify causes of low performance and to set goals and objectives for student learning. Results for the Expectations and Priorities rubric indicate that sites succeeded in making student achievement a priority and setting expectations for teachers focused on improving student learning. Only two sites, Farmville, New Mexico, and Wrightsville, Oklahoma, had means below 3.0, which indicates that school and district priorities were focused on student learning but that these priorities were not translated into expectations for teachers' work (see Expectations and Priorities Rubric in Appendix D).

Table IV.3 Descriptive Statistics for Alignment Rubric

Site	2003		2004		Change
	Mean	Std. Deviation	Mean	Std. Deviation	
Athens, LA	1.75	0.463	3.38	0.916	1.63
Bayou City, LA	2.86	0.690	2.71	0.951	-0.14
Bricktown, OK	1.56	0.527	2.22	0.833	0.67
Brookesville, OK (T)	1.14	0.378	2.00	0.000	0.86
Desert Hills, NM	2.25	0.463	2.25	0.463	0.00
Farmville, NM	1.57	0.535	2.29	0.756	0.71
Forked River, AR	1.67	0.516	2.25	0.886	0.58
Grisham, AR (T)	1.33	0.516	3.00	1.095	1.67
Highway Junction, LA	1.75	0.463	2.00	0.000	0.25
Pineland, TX (T)	2.25	0.535	3.50	0.837	1.25
Roydale, AR (T)	1.50	0.463	2.50	0.926	1.00
Wrightsville, OK	2.80	0.447	2.43	0.976	-0.37
Mean	1.85	0.691	2.53	0.886	0.68 *

T: Test site

* $p \leq .001$

Although the lowest final year mean was on the Alignment rubric, it also showed the largest mean change. Alignment of curriculum, instruction, and assessments to state standards was time-consuming work and required knowledge of standards and curricular materials. Considering that alignment constituted a major part of the work for most sites and that several states in the region were undergoing revision of standardized assessments and accountability programs during the project, it was not surprising that sites showed lower overall scores in their alignment work. All sites had means of at least 2.0 on this rubric in 2004, which suggested that they were either in the process of developing aligned curricula or had already done so. Two sites, Athens, Louisiana, and Pineland, Texas, had mean ratings higher than 3.0, indicating they were beginning to implement aligned curricula in classrooms (see Alignment Rubric in Appendix D).

Table IV.4 Descriptive Statistics for Collaboration Rubric

Site	2003		2004		Change
	Mean	Std. Deviation	Mean	Std. Deviation	
Athens, LA	2.38	1.061	3.50	1.069	1.13
Bayou City, LA	3.43	0.535	2.71	1.113	-0.71
Bricktown, OK	2.11	0.333	2.56	0.882	0.44
Brookesville, OK (T)	2.43	1.134	2.13	0.641	-0.30
Desert Hills, NM	3.00	0.926	3.38	0.744	0.38
Farmville, NM	2.29	0.951	3.14	0.900	0.86
Forked River, AR	2.83	1.329	3.13	0.835	0.29
Grisham, AR (T)	2.33	1.033	2.17	0.408	-0.17
Highway Junction, LA	3.13	0.991	3.44	0.527	0.32
Pineland, TX (T)	3.38	0.707	3.13	0.894	-0.25
Roydale, AR (T)	1.75	0.744	3.00	0.835	1.25
Wrightsville, OK	2.60	0.894	2.57	0.787	-0.03
Mean	2.63	0.990	2.92	0.897	0.29 *

T: Test site

* $p \leq .05$

The Collaboration rubric showed the most mixed results. Five sites had declines in this area, and there was the least overall growth on this rubric. In many cases, interview respondents reported that some collaboration had taken place, but these meetings appeared to be isolated events rather than ongoing work among teachers and administrators. Examples of these types of events included schoolwide or districtwide workshops to align curriculum or instruction and attempts to initiate regular collaboration among teachers in a school, grade level, or department that failed to be sustainable. Reports of these types of events tended to produce lower ratings on this rubric.

Additionally, a number of respondents discussed collaboration in the context of reporting professional development activities. Our scales specifically distinguished between collaboration for the purpose of planning and coordinating curriculum and instruction, and professional development that involved collaborative activities. In the *Working Systemically* model, these activities constituted separate competencies to be mastered by schools and districts; although

sites often integrated them in practice, it is possible that some sites may have had higher mean scores if the rubrics had integrated these two activities.

Table IV.5 Descriptive Statistics for Professional Development Rubric

Site	2003		2004		Change
	Mean	Std. Deviation	Mean	Std. Deviation	
Athens, LA	3.00	0.756	3.50	0.926	0.50
Bayou City, LA	2.86	1.069	3.71	0.756	0.86
Bricktown, OK	2.00	0.000	3.22	1.202	1.22
Brookesville, OK (T)	2.29	0.756	2.38	0.518	0.09
Desert Hills, NM	2.75	1.035	3.00	0.926	0.25
Farmville, NM	2.43	0.535	3.14	0.690	0.71
Forked River, AR	2.17	0.753	2.38	0.518	0.21
Grisham, AR (T)	2.00	0.632	3.17	0.753	1.17
Highway Junction, LA	2.50	0.756	4.00	0.000	1.50
Pineland, TX (T)	2.75	1.195	2.75	0.983	0.00
Roydale, AR (T)	3.00	1.035	3.83	0.886	0.83
Wrightsville, OK	2.00	0.707	2.00	0.577	0.00
Mean	2.51	0.861	3.09	0.950	0.58 *

T: Test site

* $p \leq .001$

Another area of strong growth was professional development. This was the only rubric in which no declines were found, although two sites, Pineland, Texas, and Wrightsville, Oklahoma, showed no change. All but four sites had a mean of at least 3.0, indicating that professional development activities in these sites were focused on building skills related to student learning needs (see Professional Development Rubric in Appendix D). The highest mean rating across all rubrics, 4.0, was found for Highway Junction, Louisiana. That rating corresponded to reports of professional development that was focused on student learning, regularly conducted, and accompanied by monitoring or follow-up.

Much of the growth on this scale can be attributed to increases in ratings of 3 and 4, coupled with decreases in ratings of 1 and 2. In 2003, half of all interviews were given a rating of 2 on this rubric. This rating indicates professional development that supported skills related to student learning but that was not intentionally focused on student learning needs of the site. In

2004, one third of interviews were given a rating of 3, which corresponds to reports of professional development that was focused on site needs. Thirty-one percent of interviews in 2004 were given a rating of 4, meaning respondents reported that professional development in their sites was focused on site needs, implemented regularly, and included some monitoring and follow-up. Most interviews that were given this rating reported use of the Professional Teaching and Learning Cycle (PTLC), which included all of these elements, in some or all parts of the site.

Working Systemically Survey Results

During the springs of 2004 and 2005, SEDL administered the *Working Systemically* survey in all 12 sites. The survey addressed alignment of standards, curriculum, instruction, and assessment; professional development; the utilization of resources toward school and district priorities; collaboration among stakeholders (school board members, administrators, teachers, families, and community); and a shared vision that all students can and are expected to achieve high academic standards. Mean scores were calculated for each survey respondent and were aggregated into mean scores for each site for both years of administration. The following table reports mean site scores for 2003–2005,²⁴ as well as mean scores and standard deviations across sites, standard deviations of the mean scores, and change scores. Possible scores ranged from 1 to 6.

²⁴ Data from 2003 were not included in longitudinal analyses due to changes in the survey instrument (see Section II for more detail), but descriptive statistics are included for examination.

Table IV.6 Descriptive Statistics for *Working Systemically* Survey, 2003–2005

Site	2003			2004			2005			Change ^a
	N	Mean	SD	N	Mean	SD	N	Mean	SD	
Athens, LA	127	4.29	0.923	49	4.87	0.634	35	4.75	0.634	-0.12
Bayou City, LA	143	4.18	0.994	133	4.18	0.777	118	4.30	0.691	0.12
Bricktown, OK	166	4.23	0.823	112	4.92	0.743	41	4.72	0.513	-0.20
Brookesville, OK (T)	38	3.25	0.921	75	3.65	0.723	63	3.80	0.858	0.15
Desert Hills, NM	47	3.79	0.971	53	3.93	0.767	51	4.20	0.830	0.27
Farmville, NM	18	3.89	0.647	45	4.00	0.867	87	4.12	0.956	0.12
Forked River, AR	39	4.34	0.939	43	4.50	0.730	63	4.34	0.845	-0.16
Grisham, AR (T)	64	4.22	0.812	52	4.61	0.720	47	4.64	0.566	0.03
Highway Junction, LA	77	4.09	1.06	93	4.57	0.855	76	4.78	0.810	0.21
Pineland, TX (T)	–	–	–	25	4.92	0.865	101	4.72	0.719	-0.40
Roydale, AR (T)	–	–	–	61	4.12	0.753	37	4.06	0.763	-0.06
Wrightsville, OK	27	4.35	0.525	31	4.08	0.438	37	4.42	0.579	0.34
Mean/Total	746	4.14	0.937	772	4.26	0.824	756	4.40	0.813	0.14 *

^a Change was calculated only from 2004 to 2005 because 2003 data were not included in longitudinal comparisons.

T: Test site

* $p \leq .001$

Mean site scores for both years fell between 4 and 5 (corresponding to Usually True and Often True on the Likert scale, respectively). Change scores were small, but the change across sites was statistically significant. This may have been due to the large sample used in the cross-site analysis. Five sites had small decreases. The remaining seven sites increased their scores, although the changes were quite small. It is important to note that sample sizes varied widely for some sites across years. In some cases, schools within sites were either added as participants (e.g., Pineland, Texas) in the project or dropped out (e.g., Bricktown, Oklahoma). In other sites, logistical issues regarding administration of the survey resulted in differences in sample sizes. The samples remained similar, however, in regard to composition and respondent characteristics (see Appendix E for descriptions of survey samples).

Changes in Systemic Work

Overall, sites showed significant increases on both measures of systemic work over time. However, only the interview score growth appeared to be meaningful. The following graphs display the slopes of changes found in the mean scores across sites for each measure.

Figure IV.1 Increase in Interview Scores, 2003–2004

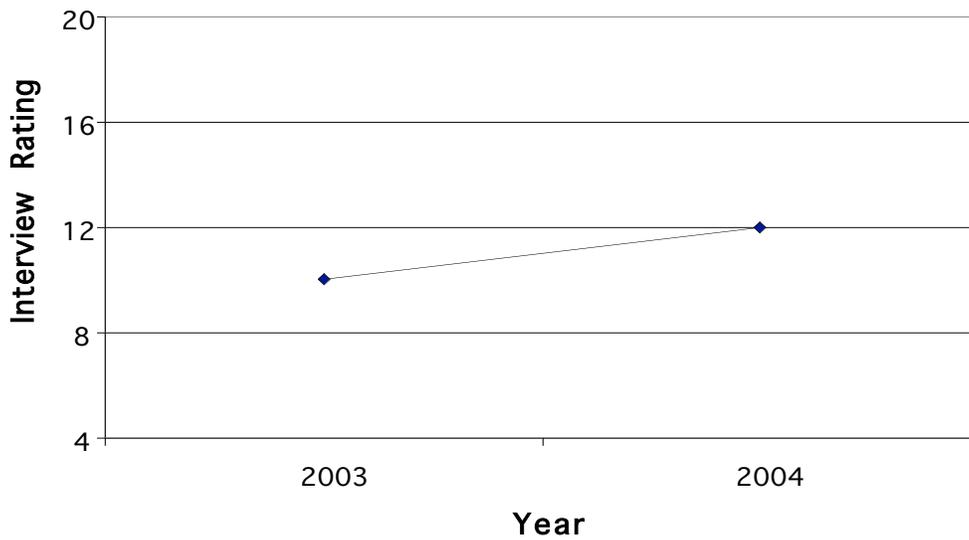
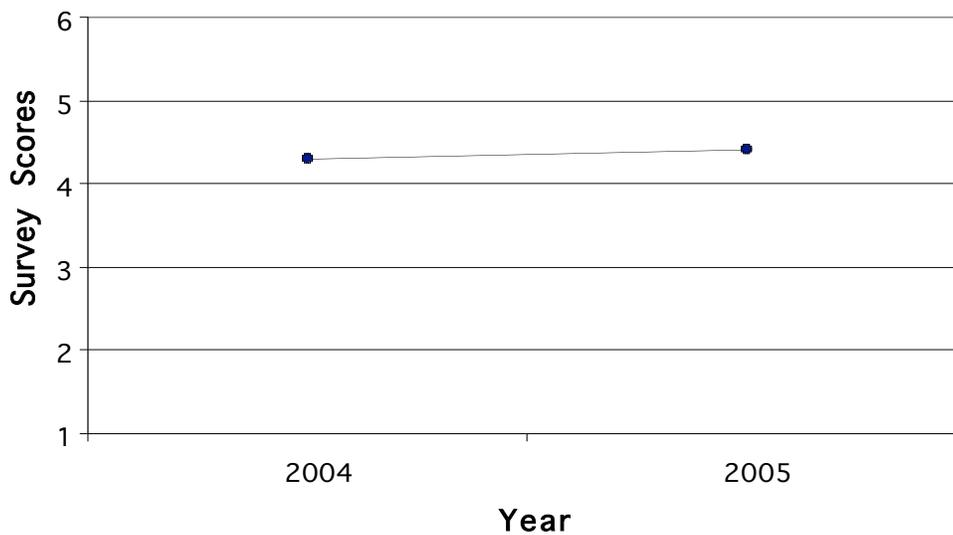


Figure IV.2 Increase in *Working Systemically* Survey Scores, 2004–2005



The rate of change for the interview scores was 12.1%, while the rate of change for the survey scores was only 1.8%. This could have been due to differences in the sampling strategies employed for each measure. Entire faculties of schools in which SEDL worked took the survey, as well as members of groups and committees involved in implementing the *Working Systemically* model (for example, district leadership teams). This strategy allowed data collection with a large, heterogeneous group that could provide broad information on instructional and organizational practices in the districts. Interviews, on the other hand, occurred with a more select group of individuals with specific responsibility for working with field staff in the sites. This group was knowledgeable about the goals, plans, and activities of those involved in conducting the work with SEDL. They were able to discuss district practices at some length. Survey respondents, who were predominately teachers, may not have had such thorough knowledge of plans and practices that were in development and thus may not have considered them when reporting their perceptions of their schools and districts. While these new structures and activities may not yet have been fully implemented or affected the daily experiences and perceptions of most classroom teachers (thus the relative lack of change in survey results), they seem to have influenced the work and attitudes of school and district leaders. In other words, the interviews most likely measured the intensity with which this smaller group approached the work, while the survey results hint at a modest spread of the model into regular instructional and organizational practices.

Differences in Development of Systemic Work

The research team conducted further analyses of the interview and survey data to elaborate this apparent distinction between the intensity and extent of sites' development of systemic work. The team analyzed differences between all sites, test and development sites, and teachers and administrators. Analysis of variance (ANOVA) methods were used to determine the significance of any differences.

Comparisons of Sites

This analysis was intended to determine whether some sites differed from others in their development over time of systemic work. Univariate ANOVAs were conducted using site and year as independent variables to test whether sites differed over time.

Interview results.

As reported in Table IV.7 below, the ANOVA resulted in significant effects for site, year, and the interaction between site and year. These results indicate that there was variation in sites' change in working systemically as measured in the interviews.

Table IV.7 Univariate ANOVA on Interview Scores by Site and Year

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	458.972	23	19.955	4.553	.000	.405
Intercept	21122.689	1	21122.689	4819.544	.000	.969
Site	173.269	11	15.752	3.594	.000	.204
Year	162.829	1	162.829	37.152	.000	.194
Site * Year	110.256	11	10.023	2.287	.013	.140
Error	674.938	154	4.383			
Total	22760.000	178				
Corrected Total	1133.910	177				

$R^2 = .405$ (Adjusted $R^2 = .316$)

ANOVAs were also conducted on each rubric (Expectations and Priorities, Alignment, Collaboration, and Professional Development). Site mean ratings for 2003 and 2004 were significantly correlated for Expectations and Priorities and Professional Development. Levene's test of equality of error variances indicated that the error scores (the differences among the scores in each group used in the analysis) were not equal.²⁵ Additionally, site mean scores on both rubrics were strongly correlated. Due to these problems, the ANOVA results were

²⁵ For Expectations and Priorities, Levene's $F = 1.591, p = .009$. For Professional Development, Levene's $F = 2.375, p = .001$.

unreliable (Glass & Hopkins, 1996), and are not included in presentations and interpretations of findings in this report.

These data problems did not occur for the two remaining rubrics, Alignment and Collaboration. Results of univariate ANOVAs on these two rubrics are presented in Table IV.8 and Table IV.9.

Table IV.8 Univariate ANOVA on Alignment Scores by Site and Year

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	64.499	23	2.804	6.387	.000	.488
Intercept	847.861	1	847.861	1930.962	.000	.926
Site	28.063	11	2.551	5.810	.000	.293
Year	19.833	1	19.833	45.169	.000	.227
Site * Year	16.524	11	1.502	3.421	.000	.196
Error	67.619	154	.439			
Total	991.000	178				
Corrected Total	132.118	177				

$R^2 = .488$ (Adjusted $R^2 = .412$)

Table IV.9 Univariate ANOVA on Collaboration Scores by Site and Year

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	45.440	23	1.976	2.645	.000	.283
Intercept	1336.814	1	1336.814	1789.928	.000	.921
Site	26.122	11	2.375	3.180	.001	.185
Year	3.096	1	3.096	4.145	.043	.026
Site * Year	14.144	11	1.286	1.722	.073	.110
Error	115.015	154	.747			
Total	1537.000	178				
Corrected Total	160.455	177				

$R^2 = .283$ (Adjusted $R^2 = .176$)

Both analyses resulted in significant main effects for site and year. The analysis on Alignment scores also found a significant interaction effect between site and year. The effects of

site and year on Alignment were also substantively strong. The model accounted for over 40% of the variance in Alignment scores (adjusted R squared = 0.412), and eta-squared statistics indicated that almost one third of the variance (0.293) could be accounted for by site.

Collaboration scores were weaker, although the independent variable of site similarly accounted for the largest proportion of variance in scores. However, the lack of a significant interaction effect between site and year indicates that while scores varied by site in each year, and cross-site scores varied each year, sites were not different in their changes over time. This can be seen by examining the change scores in Table IV.4, which have a more limited range than those found for Alignment (shown in table IV.3).

Working Systemically survey results.

The results for *Working Systemically* survey data proved problematic for reasons that mirrored those for analyses of the Expectations and Priorities and Professional Development interview rubrics. Levene's test of equality of error variances indicated that the error scores were not equal (Levene's $F = 2.826$; $p < .000$). Survey sample sizes also varied widely for some sites across years, as reported in Table IV.6. Additionally, site mean scores for 2004 and 2005 were highly correlated (Pearson's $r = .864$, $p < .000$). Due to these problems, the ANOVA results were unreliable and are not included in presentations and interpretations of findings in this report.²⁶

Since the survey data proved difficult to interpret in these analyses, interpretations must be made on only the interview results. Those comparisons revealed considerable variation in sites' progress in working systemically. The context of the site and the activities undertaken as

²⁶ With survey scores as the dependent variable, Year $F = 4.670$, $p = .031$; Site $F = 22.495$, $p = .000$; Interaction $F = 1.606$, $p = .091$. Levene's test was also significant when interview sum scores were used as the dependent variable (Levene's $F = 1.976$, $p = .008$). However, sample sizes within sites were relatively stable across years, and site scores were not correlated (Pearson's $r = .171$, $p = .595$). Because two-factor ANOVAs are robust to inequality of error variances when samples are balanced and when factors are independent (Glass & Hopkins, 1996), results of that analysis are included in this report.

part of their work with the model appear to have influenced the extent of their systemic work. However, patterns of differences were difficult to determine. The variation found was likely due to idiosyncratic differences in the ways sites were working with field staff to incorporate the model into their practices, differences in subject matter and grade levels being addressed, and other contextual issues. Some of these issues will be addressed in the exploratory site analyses.

Comparisons of Test and Development Sites

Next, the research team compared test and development sites. The following tables provide descriptive statistics and results of the univariate ANOVA comparing test and development sites on interview sum scores.

Table IV.10 Descriptive Statistics for Test/Development Interview Scores

Interview Year	Site Group	Mean	Std. Deviation	N
<i>2003</i>	Development	10.22	2.153	58
	Test	9.62	2.665	29
	<i>Total</i>	10.02	2.338	87
<i>2004</i>	Development	12.02	2.311	63
	Test	11.89	2.455	28
	<i>Total</i>	11.98	2.343	91
<i>Total</i>	Development	11.16	2.401	121
	Test	10.74	2.787	57
	<i>Total</i>	11.02	2.531	178

Table IV.11 Univariate ANOVA on Interview Scores by Test/Development and Year

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	177.334	3	59.111	10.752	.000	.156
Intercept	18530.124	1	18530.124	3370.605	.000	.951
Year	159.859	1	159.859	29.078	.000	.143
Test/Development	5.108	1	5.108	.929	.336	.005
Test/Development * Year	2.234	1	2.234	.406	.525	.002
Error	956.576	174	5.498			
Total	22760.000	178				
Corrected Total	1133.910	177				

$R^2 = .156$ (Adjusted $R^2 = .142$)

The ANOVA resulted in a nonsignificant main effect for test/development status and a nonsignificant interaction between test/development status and year. Interview scores between the two groups were similar over time. There was a significant, though small, main effect for year indicating positive change over time.

A univariate ANOVA was also conducted comparing test and development sites on *Working Systemically* survey scores. Descriptive statistics are reported in Table IV.12 below. Results of the univariate ANOVA are shown in Table IV.13.

Table IV.12 Descriptive Statistics for Test/Development Working Systemically Survey Scores

Survey Year	Site Group	Mean	Std. Deviation	N
2004	Development	4.30	.802	559
	Test	4.17	.875	213
	Total	4.26	.824	772
2005	Development	4.41	.802	508
	Test	4.37	.837	248
	Total	4.40	.813	756
Total	Development	4.35	.803	1,067
	Test	4.28	.860	461
	Total	4.33	.821	1,528

Table IV.13 Univariate ANOVA on *Working Systemically* Survey Scores by Test/Development and Year

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	9.441	3	3.147	4.702	.003	.009
Intercept	23823.689	1	23823.689	35593.126	.000	.959
Year	7.766	1	7.766	11.603	.001	.008
Test/Development	2.117	1	2.117	3.163	.076	.002
Test/Development * Year	.802	1	.802	1.198	.274	.001
Error	1020.065	1524	.669			
Total	29654.406	1528				
Corrected Total	1029.506	1527				

$R^2 = .009$ (Adjusted $R^2 = .007$)

The main effect for test/development status was not significant, and neither was the interaction between test/development and year. There was a significant, though small, main effect for year. These results suggest that changes in the survey scores over time were similar between test and development sites. The year effect was highly significant, indicating, as in previous analyses, a positive change over time.

Taken together, these analyses indicate that test and development sites did not differ significantly in working systemically. Of course, test sites became involved with the *Working Systemically* model later in the project than most development sites and were using a more refined and developed version of the *Working Systemically* model. The fact that these sites showed results similar to development sites suggests that they developed practices and goals similar to those of development sites in a shorter period of time.

Comparisons of Teachers and Administrators

Analyses of changes in systemic work over time across sites showed higher growth on interview scores than on survey scores. Although the instruments differed in both substance and methodology, they were highly correlated, suggesting the differences in results may have been due to differences in the samples on which data were collected. The *Working Systemically* survey

samples contained predominately teachers, while the interview samples had higher proportions of administrators. As discussed in Section III, field staff strategies emphasized work with school and district administrators who spearheaded the work in their sites. These distinctions between teachers and administrators prompted the research team to conduct analyses comparing these groups on their development of systemic work.²⁷

Interview results.

Univariate ANOVAs were conducted on both measures using role and year as independent variables. Table IV.14 reports descriptive statistics for interview sum scores. Table IV.15 reports results from the ANOVA on interview sum scores.

Table IV.14 Descriptive Statistics for Interview Scores by Role

Interview Year	Group	Mean	Std. Deviation	N
2003	Teachers	9.89	2.334	38
	Administrators	10.19	2.340	48
	<i>Total</i>	10.06	2.328	86
2004	Teachers	11.40	2.444	30
	Administrators	12.52	2.261	50
	<i>Total</i>	12.10	2.379	80
<i>Total</i>	Teachers	10.56	2.482	68
	Administrators	11.38	2.570	98
	<i>Total</i>	11.04	2.560	166

²⁷ Respondents in other roles, such as counselors, teacher aides, district support staff, and parents were also included in interview and survey samples. However, the sizes of these groups were very small, and these roles were not systematically included in samples across years. Therefore, these roles were not included in the longitudinal analyses presented here.

Table IV.15 Univariate ANOVA on Interview Scores by Role and Year

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	198.133	3	66.044	12.123	.000	.183
Intercept	19269.004	1	19269.004	3536.913	.000	.956
Year	146.577	1	146.577	26.905	.000	.142
Role	19.863	1	19.863	3.646	.058	.022
Year * Role	6.810	1	6.810	1.250	.265	.008
Error	892.571	162	5.448			
Total	21321.000	166				
Corrected Total	1080.705	165				

$R^2 = .183$ (Adjusted $R^2 = .168$)

Again, the main effect for year was significant, indicating overall growth over time. The effect size was small, however. The main effect of role was not significant, and neither was the interaction between year and role. This indicates that teachers and administrators did not differ significantly in their interview sum scores over time. However, the effect of role approached significance, so analyses were conducted on the mean scores for each interview rubric to determine whether there were any differences on the more specified rubrics that may not have been apparent in the comparison of sum scores. Significant results were obtained for the rubrics of Expectations and Priorities and Professional Development. In both cases, administrators' scores were higher than those of teachers. The following tables provide mean interview sum scores and sample sizes for teachers and administrators on all interview rubrics.

Table IV.16 Descriptive Statistics for Expectations and Priorities Scores by Role

Interview Year	Role	Mean	Std. Deviation	N
2003	Teachers	3.05	1.012	38
	Administrators	3.02	.812	48
	Total	3.03	.900	86
2004	Teachers	3.07	.828	30
	Administrators	3.70	.789	50
	Total	3.46	.856	80
Total	Teachers	3.06	.929	68
	Administrators	3.37	.866	98
	Total	3.24	.902	166

Table IV.17 Descriptive Statistics for Alignment Scores by Role

Interview Year	Role	Mean	Std. Deviation	N
2003	Teachers	1.82	.652	38
	Administrators	1.90	.722	48
	<i>Total</i>	1.86	.689	86
2004	Teachers	2.53	.860	30
	Administrators	2.60	.926	50
	<i>Total</i>	2.58	.897	80
<i>Total</i>	Teachers	2.13	.827	68
	Administrators	2.26	.900	98
	<i>Total</i>	2.20	.871	166

Table IV.18 Descriptive Statistics for Collaboration Scores by Role

Interview Year	Role	Mean	Std. Deviation	N
2003	Teachers	2.66	.966	38
	Administrators	2.65	1.000	48
	<i>Total</i>	2.65	.979	86
2004	Teachers	2.97	1.033	30
	Administrators	2.98	.845	50
	<i>Total</i>	2.98	.914	80
<i>Total</i>	Teachers	2.79	1.001	68
	Administrators	2.82	.934	98
	<i>Total</i>	2.81	.959	166

Table IV.19 Descriptive Statistics for Professional Development Scores by Role

Interview Year	Role	Mean	Std. Deviation	N
2003	Teachers	2.37	.751	38
	Administrators	2.63	.937	48
	<i>Total</i>	2.51	.864	86
2004	Teachers	2.83	.874	30
	Administrators	3.24	1.001	50
	<i>Total</i>	3.09	9.70	80
<i>Total</i>	Teachers	2.57	.834	68
	Administrators	2.94	1.013	98
	<i>Total</i>	2.79	.959	166

Univariate ANOVAs were conducted on all measures using role and year as independent variables. Results for the Alignment and Collaboration scales were not significant for either main effects of role or for the interactions between role and year. Main effects for year were significant but small.²⁸ Results for the Expectations and Priorities and Professional Development scales are presented in Table IV.20 and Table IV.21.

Table IV.20 Univariate ANOVA on Expectations and Priorities Scores by Role and Year

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	15.121	3	5.040	6.848	.000	.113
Intercept	1640.774	1	1640.774	2229.152	.000	.932
Year	4.782	1	4.782	6.497	.012	.039
Role	3.601	1	3.601	4.892	.028	.029
Year * Role	4.403	1	4.403	5.982	.016	.036
Error	119.241	162	.736			
Total	1878.000	166				
Corrected Total	134.361	165				

$R^2 = .113$ (Adjusted $R^2 = .096$)

Table IV.21 Univariate ANOVA on Professional Development Scores by Role and Year

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial η^2
Corrected Model	18.242	3	6.081	7.385	.000	.120
Intercept	1218.850	1	1218.850	1480.398	.000	.901
Year	11.606	1	11.606	14.097	.000	.080
Role	4.378	1	4.378	5.317	.022	.032
Year * Role	.224	1	.224	.272	.603	.002
Error	133.379	162	.823			
Total	1443.000	166				
Corrected Total	151.620	165				

$R^2 = .120$ (Adjusted $R^2 = .104$)

²⁸ Alignment: Role $F = .335$, $p = .564$; Year $F = 31.438$, $p < .000$; Interaction $F = .003$, $p = .958$.
 Collaboration: Role $F = .000$, $p = .997$; Year $F = 4.519$, $p = .035$; Interaction $F = .007$, $p = .933$

Both analyses found significant main effects for year and role. The interaction between year and role was also significant for Expectations and Priorities. Examination of the mean scores in Tables IV.16 shows that teachers and administrators had nearly identical means on this scale in 2003, but those means diverged in 2004, with administrators having a greater mean score. On the Professional Development scale (see Table IV.19), administrators had higher means in both years, but both groups increased their scores on this scale over time, resulting in a nonsignificant interaction between year and role.

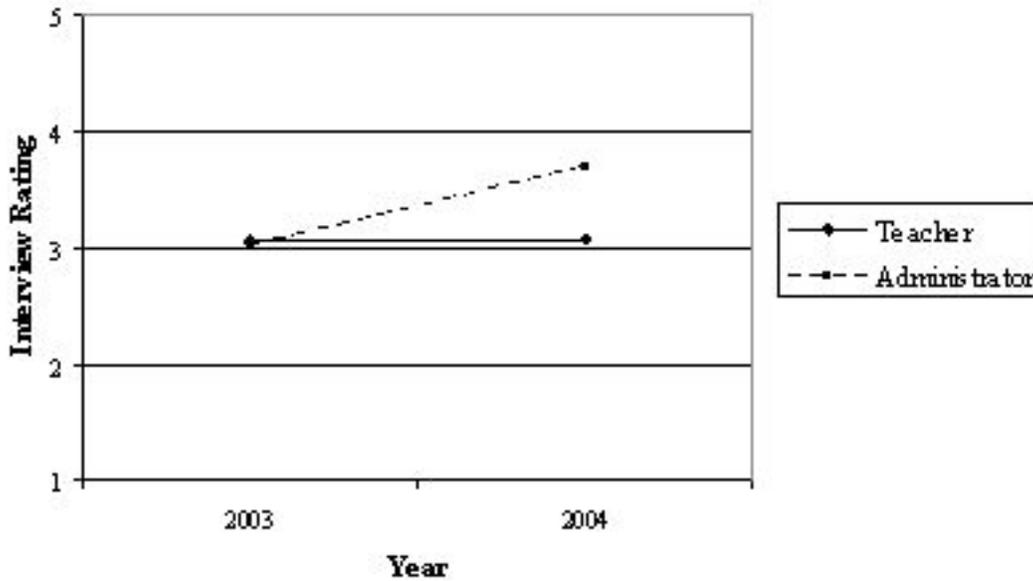
Working Systemically survey results.

Comparisons of teachers and administrators on survey scores were problematic due to inconsistencies in the data. Levene's test of equality of error variances for these data was significant ($F = 4.454, p = .004$). Sample sizes for these groups were also very different each year, making ANOVA results unreliable. In any case, neither the main effect for role nor the interaction between role and year were statistically significant (role: $F = 3.082, p = .079$; interaction: $F = .260, p = .610$). The main effect of year was not significant either ($F = 3.803, p = .051$).

In general, interviews with administrators yielded higher scores than those of teachers. This could reflect that administrators (who were predominately principals and district administrators) had more knowledge about the extent to which practices were being implemented in other schools and grade levels in their districts. Teachers, on the other hand, were often able to comment only on activities within their own buildings, grade levels, or departments. Therefore, differences between teachers and administrators might simply reflect the different realms in which these groups tended to work. On the other hand, the full range of ratings (1–5) were assigned to both groups, indicating that some teachers did have knowledge of practices taking place outside their particular spheres of work.

Another possible explanation is that administrators may have been more optimistic about the effects of new practices and initiatives than teachers who were engaged in daily classroom instruction. The results for Expectations and Priorities are particularly suggestive of this. The following graph illustrates the slopes of change on this scale between teachers and administrators.

Figure IV.3 Change in Expectations and Priorities Scale by Role



The two groups had virtually identical means in 2003, but the mean for administrators rose by nearly one point in 2004 while the mean for teachers remained almost the same. Given the emphasis in this scale on the communication of student learning goals and their translation into practice, this finding is troubling. It indicates that the school and district leaders interviewed perceived themselves as improving their emphasis on student achievement in both rhetoric and practice, but that the teachers interviewed saw little change. The teachers' mean of about 3.0 in both years corresponds in the scale to consistency of messages about school and district priorities and expectations of teachers focused on student learning. Higher ratings indicate that more specific requirements of teachers to engage in activities or processes to improve student learning reinforced these messages. The higher scores of administrators in 2004 indicates that they

regarded their new practices and initiatives to be progressing to a greater degree than did the teachers to whom those initiatives were directed.

Exploratory Analyses of Selected Sites

Using two measures of working systemically (interviews and surveys) enabled the research team to compare sites across measures quantitatively. Analyses of differences between sites over time indicated that there was variation in development of systemic work, but the reasons behind these differences were unclear. In addition, it was apparent that there were some sites that scored consistently high on measures of working systemically, others that scored consistently low, and some that showed mixed results across interview and survey measures. To illuminate the ways in which sites developed systemic work and possibly explain variations in outcomes among sites, the research team conducted qualitative analyses of four sites that had either high and low scores on interview and survey measures relative to other sites and those that had mixed results on the two measures.

These analyses were intended to explicate the results of the quantitative analyses. Examination of sites with high and low rankings on both measures is helpful in offering insight as to the particular practices and activities that were important in developing systemic work.²⁹ Given the strong correlation between the *Working Systemically* survey and the interview scores, mixed results indicate unusual patterns of results that could prove insightful in examining the *Working Systemically* model (Erzberger & Kelle, 2003).

To select sites, rankings were developed for final year means on each measure and for change scores on each measure. Table IV.22 reports the rankings. The selected sites are in bold.

²⁹ Since the focus of these exploratory analyses was to illuminate findings, explain differences between results on the two measures, and generate possible explanations of quantitative findings that might pertain to all sites, only interview and survey data were used. These data were also used to develop the quantitative rankings by which sites were selected, making them especially pertinent. While data from site contact records may have provided richer detail about the contexts, histories, and nature of reform work in these sites (depending on the extent and nature of SCR data available), inclusion of those data would have been beyond the scope and intention of these analyses.

Table IV.22 Rankings of Sites by Final Year Mean and Change Scores

Site	Final Year		Change	
	2004 Interview Mean Rank	2005 Survey Mean Rank	Interview Change Rank	Survey Change Rank
Athens, LA	1^a	2^a	1^a	10^b
Bayou City, LA	3	8	11	7
Bricktown, OK	8	5	4	1 ^a
Brookesville, OK (T)	11^b	12^b	8	5
Desert Hills, NM	6	9	10	3
Farmville, NM	9	10	6	6
Forked River, AR	10	7	7	11 ^b
Grisham, AR (T)	6	4	3	8
Highway Junction, LA	4	1 ^a	5	4
Pineland, TX (T)	5	3 ^a	9	12 ^b
Roydale, AR (T)	2	11^b	2^a	9
Wrightsville, OK	12^b	6	12^b	2^a

T: Test site

^a Mean was more than one standard deviation above the mean for all sites

^b Mean was more than one standard deviation below the mean for all sites

Rankings were examined first in terms of their distance from the means across sites, then in terms of their overall patterns across measures. The four sites selected were Athens, LA, Brookesville, Oklahoma, Roydale, Arkansas, and Wrightsville, Oklahoma. As reported in the table, Athens had consistently high scores, with the exception of their change score for the survey. However, the site had the second highest mean score on the survey in 2005. Brookesville, on the other hand, had low scores and rankings on both measures, and their change scores were middle range.

Roydale and Wrightsville had more unusual patterns of results. Roydale ranked second to last on the 2005 survey but second from the top in the 2004 interview scores. They also had high growth in interview rankings but low growth in survey results. Wrightsville had the lowest mean scores on both the 2004 interviews and the least change in interview scores. However, the site scored higher on the final year survey and had the second highest growth on the survey. Given

these unusual patterns of results, the research team conducted exploratory analyses of Roydale and Wrightsville to provide some possible explanations for their results.

Athens, Louisiana

Athens was one of the top performers on both measures of working systemically. This site had high scores on both interviews and the survey, as well as the greatest change in interview scores. Although Athens had the second highest score on the *Working Systemically* survey, the site showed relatively little change from 2004 to 2005. Table IV.1 shows that Athens had the second highest mean survey score in 2004 as well. Given the relatively high scores, there may have been a ceiling effect curtailing growth on the survey for this site.

Examination of survey results by item showed significant decreases on several items related to provision and use of resources. These items included the following: supplies essential for instruction are available in sufficient quantity; administrators ensure that teachers have necessary resources to deliver high-quality instruction; and the school budget directly supports the school's priorities. Interview data did not produce any findings relevant to availability or use of resources, although the superintendent noted in a 2003 interview that the district had been unable to purchase science textbooks the previous year due to budget shortfalls. Such financial difficulties could have negatively influenced survey respondents' perceptions about provision of resources in the district.

Athens showed the most growth in interview scores. Tables IV.3–IV.6 show that this site had gains of over one point on the Expectations and Priorities, Alignment, and Collaboration scales. It showed growth in Professional Development as well, although not as strong. The site's increase appeared to derive from greater coherence of practices and more focus on student achievement that was apparent across ratings. Interview respondents in 2004 frequently discussed increased collaboration and work on curriculum alignment taking place in the district,

describing an elaborate meeting structure in some cases. For example, one middle school teacher noted,

[W]e have four eighth grade teams, and three seventh grade teams. We have a team leader on each of those teams and they meet with administration every week and they come back to us and meet with us each day. [T]hen we have faculty-staff development meetings. We also have every other Thursday, the students get off a little early and we have a couple of hours in the afternoon where we meet for staff development within our curriculum both seventh and eighth grade teachers. And each day we meet as a team and we meet in a team room. Each team has an hour to meet there. The activities during this time vary depending on what needs to be done, meeting parents, curriculum [alignment], discuss[ing] student issues with students.

Embedded within this description were practices related to collaboration, alignment, and professional development, as well as parent involvement and student learning. Additionally, the structures incorporated vertical and horizontal alignment and planning within and across grade levels. Respondents also mentioned expectations for teachers that were quite specific in some cases, including differentiating instruction, documentation of lessons aligned to standards, and use of benchmark assessments. The district curriculum supervisor noted that the district expected teachers

to understand the needs of the students that they teach, [t]o know the what of what they are going to teach, [t]o use research based strategies to do the how and to assess in a method similar to how students will be tested that will generate a means for them to zero in on student weaknesses and intervene with students who don't master the skills that we need mastered.

Taken together, these comments painted a picture of a district becoming more focused on student achievement and instructional improvement and implementing practices to achieve those goals.

Brookesville, Oklahoma

Brookesville was one of the low performers on both measures of working systemically. The site had low scores on both interview scores and the *Working Systemically* survey. Brookesville ranked higher on change scores for both measures, although not greater than one standard deviation above the mean for all sites.

Brookesville had a significant increase on the Alignment rubric, although the site's mean rating (2.00) was still low. In examining interview responses, it appeared that the district had begun work on alignment but was still in the beginning stages of the process. In 2003, respondents noted that teachers typically taught from their textbooks and developed individual curriculums for their classrooms. A high school principal, for example, said,

Before SEDL, each discipline had its own way of teaching. [Teachers] weren't teaching to a test, [they] didn't have any type of alignment, so no one knew what the others were teaching. So there really wasn't any kind of alignment. We were taking a shot in the dark and teaching from the textbooks and hoping that, because they were state adopted, that they were aligned with the standards and were covering the material.

In 2004, respondents discussed steps being taken in the district to align curriculum to standards, such as requiring teachers to specify the standards being covered in their lesson plans. These steps still warranted a relatively low rating on the Alignment rubric but nonetheless demonstrated improvement.

Survey data also provided some evidence of improvement in alignment. Of three items on which there were significant increases from 2004 to 2005 in Brookesville, one dealt with alignment ("Our school's curriculum is closely aligned with the state's standards"). The other two were related to professional development ("Research-based instructional strategies are the focus of our professional development") and internal accountability ("The individuals responsible for implementing each part of the school plan are held accountable for seeing that those plans and objectives are carried out").

However, these improvements in alignment were offset by declines in ratings for collaboration. Few respondents reported any kind of formal collaboration between teachers for planning or discussing curriculum and instruction. Teachers in this site seemed to teach primarily in isolation from one another and resisted working together, although there appeared to be a growing acknowledgement of the importance of collaboration. A reading specialist commented,

It [collaboration] happens some and I think it's starting to happen more, but I think that somehow teachers have this fear of people knowing what's going on or something so we still have some locked doors. We can't just shut our doors and ignore everybody else. We do need to work together.

In all, Brookesville apparently began to make progress in developing practices that would enable the site to work systemically. Given more time, the progress may have continued.

Roydale, Arkansas

Roydale's results showed high interview scores both in the final year mean and change from 2003 to 2004. However, the site had both a low final year mean score and a low change score on the *Working Systemically* survey.

The increases in interview scores came about primarily from increases in the alignment and collaboration ratings. In fact, Roydale had the highest change on the collaboration rubric. Interview responses indicated the district began to implement strategies to align curriculum to standards, such as developing pacing guides that identified standards to be taught in 9-week periods, requiring teachers to document alignment of their lesson plans to specific standards, and mapping grade-level curriculums to standards. Respondents discussed development of common preparation periods for teachers during the school day intended to facilitate collaboration. One school principal also noted the creation of regular collaborative meetings for faculty after school.

Results on the *Working Systemically* survey for Roydale were relatively stable over time. Only two items had significant changes. One item concerning facilities ("Space exists in our school building for staff and others to work together") had a significant increase. Another item

related to internal accountability (“Administrators hold teachers accountable for student learning”) had a significant decrease.

Considering this relative stability, the mixed rankings in this site may have been due to positive developments in alignment and collaboration in some parts of the site that were relevant to interview respondents (all respondents but one in 2004 were administrators) but had not become integrated well enough into the regular practices of the district to influence respondents to the survey, who were primarily teachers.

Wrightsville, Oklahoma

Wrightsville was also selected because its results were mixed between the two measures. This site had the only decrease over time in mean interview scores, yet had the second highest score on the *Working Systemically* survey in 2005.

Examination of results for individual interview rubrics (see Tables IV.3–IV.6) showed that the site’s decrease was primarily due to decreases in the Expectations and Priorities rubric. The interview data suggested that in 2003 the leaders of Wrightsville’s work with SEDL had a strong focus on using student achievement data and assessments to make decisions about curriculum and instructional planning. Reports of teacher expectations also centered on their use of student assessments to individualize instruction.

However, this strong focus on data-driven instruction and decision making was not as apparent in the 2004 interview data. By that time, some administrators may have become preoccupied with other issues, such as facilities and safety. For example, in 2003 the district superintendent responded to a question about priorities in the district by focusing on teachers’ use of data in instruction:

We are trying to get [our teachers] to study data from tests; to identify weaknesses and focus on those as the priority in the classroom. Once the kids master those areas, then [teachers can] pick other areas where there are weaknesses and work on them. Our greatest potential for growth comes from [our] weakest areas.

In response to a similar question in 2004, the same superintendent talked primarily about safety concerns in the district:

Priorities for [elementary school], first of all, would be safety for students, education for students, and trying to improve the academic achievement of those students. Security is an issue at [elementary school] because we are dealing with a little bit more volatile population. We work really hard to maintain that safe environment for our teachers and for our students and for the whole staff there.

In a later question in the same interview, the superintendent again pointed to security as being a predominant issue in the district. Other respondents who were interviewed in both years had similar changes in focus away from the use of student data in decision-making.

However, survey results for the same school year (2004–2005) suggested that teachers and other staff in the district may not have felt the effects of this dilution of focus. Wrightsville had the second highest increase in survey scores and had significant increases on a number of items. Table IV.23 reports these items.

Table IV.23 Significant Changes on *Working Systemically* Survey Items–Wrightsville, OK

Survey Item	2004 Mean	2005 Mean	Change
Space exists in our school building for staff and others to work together.	2.58	3.57	0.99 **
Administrators know what is to be taught and learned in each grade and subject area.	3.61	4.49	0.88 **
Our school provides sufficient time for professional development.	3.48	4.35	0.87 ***
Teachers have a voice in how resources are allocated.	3.06	3.81	0.75 *
Our school has the necessary resources to support research-based practices that promote student learning.	3.43	4.16	0.73 **
Data are used to make decisions in our district and schools.	3.94	4.65	0.71 **
Resources are reallocated based on reviews of student needs.	3.27	3.95	0.68 **
Administrators' actions demonstrate their commitment to improving student learning.	4.33	4.95	0.62 *
Administrators and teachers work together to develop a shared vision of how the school should best meet the needs of all students.	3.87	4.49	0.62 *
Our professional development programs match our school's priorities.	3.63	4.22	0.59 *
Professional development programs improve overall school performance.	3.83	4.41	0.58 *
The purpose of professional development programs is to improve student learning.	4.32	4.89	0.57 *
Administrators hold teachers accountable for student learning.	4.52	5.08	0.56 *
Research-based instructional strategies are the focus of our professional development.	4.06	4.62	0.56 *
Our school's curriculum is closely aligned with the state standards.	5.47	4.81	-0.66 **

* $p \leq .05$
 ** $p \leq .01$
 *** $p \leq .001$

With the exception of the item on alignment (which mirrored findings from the Alignment interview rating scale), most of these items related to quality of professional development and confidence in leadership. The professional development items pertained to the focus of activities on student learning and strategies to improve student achievement. The leadership items dealt with collaboration between administrators and teachers, development of

vision, and internal accountability. These results suggest that staff in the schools continued to feel satisfied with their professional development and confident in the goals and commitment of school and district leadership.

Upon closer examination, however, it appeared that much of Wrightsville's increase in survey scores was due to perceived improvements in resources and leadership. Of the 14 items for which significant increases were found, five related to resources (which here included time and space as well as funds and materials). The greatest change was found for an item concerning space for collaboration, which primarily measured adequacy of facilities. The other items relating to resources addressed allocation and adequacy of resources. While these areas were important to systemic work as theorized in the model, they did not directly relate to the capacities measured in the interview ratings. Similarly, the items relating to leadership, which focused on knowledge, vision, and accountability, did not directly map onto the interview rubrics. These differences in instruments and concepts measured could account for the overall differential results obtained for Wrightsville.

Still, the differing results for measures of professional development did indicate variations within the site on practices in this area. Judging from the survey results, teachers in Wrightsville perceived the site's professional development program to be focused on student learning and school priorities. However, the site had a mean of 2.0 on the Professional Development interview rubric in both 2003 and 2004, which corresponded to reports of professional development that built some skills related to student learning but were not intentionally focused on needs in the site. It is possible that substantial improvements were made in this area between Fall 2004 (when interviews were conducted) and Spring 2005 (when surveys were administered). However, the professional development activities described by interview respondents had little focus or coherence. Most described the professional development program as consisting of workshops and seminars that teachers selected individually based more on their needs and preferences than on a systematic analysis of instructional needs throughout the site or

its schools. The only exception was the responses of one teacher who was involved in the site's pilot of the PTLTC, which took place in only one grade level at one school. This type of professional development program may have appeared focused to teachers using the more norm-referenced survey instrument but did not demonstrate the kind of coherence and thoroughness measured in the more criterion-referenced interview rubrics.

Discussion

Some sites engaged in the *Working Systemically* model improved considerably in working systemically. That is, they developed skills and practices that enabled them to better integrate and direct the various facets of their school systems toward achieving student learning goals. Others did so to a lesser degree. Results of site comparisons varied, but the finding of growth across all sites over time was consistent.

It is important to note that the magnitude of these effects was quite small. Consistent growth over time was found on the *Working Systemically* survey, for example, but sites overall increased their scores by less than one tenth of a point. Interview results showed more growth, but changes on individual rubrics were modest at less than 1 point on a 5-point scale. Additionally, the amount of variation in scores explained by the variables examined in comparative analyses (year, test or development status, role of respondent) was small even when effects were statistically significant. Although sites increased their systemic work, their growth was relatively limited.

One explanation for this finding may be that the instruments used to measure systemic work were designed to be comprehensive in nature and thus may have obscured some more notable changes. This is particularly true of the *Working Systemically* survey, which consisted of 50 items that were all intercorrelated. In essence, the instrument was equivalent to a single index of systemic work measured by 50 items. Such a lack of categorizing leads to loss of information about variation and change and less sensitive analyses (Pedhazur, 1997).

There are also more substantive explanations for the limited growth found here. First, the changes in school and district practices promoted in the *Working Systemically* model, such as curriculum alignment and improvement of instruction, can be complex and difficult to implement (Corcoran & Christman, 2002). The restricted time period of our analyses may have made more substantial growth difficult to detect. The time period of the project itself may have also made this difficult. Only six of the 12 sites studied participated in SEDL's work for more than 3 years. Two sites participated for only 2 years. The types of substantial organizational, curricular, and instructional reforms sought in this project may take much longer to realize.

Second, contextual variations in the sites over the course of the project likely influenced outcomes. As discussed in Section I, the stages of the model did not follow a linear path but accommodated more realistic iterative and cyclical patterns. Thus, progress in sites sometimes slowed down temporarily as educators and field staff reformulated plans or objectives in response to new information. Accordingly, Table IV.1 shows that some sites experienced drops in systemic work before improving. Others increased, then declined. Still others experienced steady but slow growth. This indicates that changes in strategies and practices associated with the model, as well as other changes taking place in the sites such as leadership turnover and introduction of new initiatives, affected their progress toward systemic work in ways that were not always linear and did not always lead to growth on these measures. Leadership turnover may have been an important issue because 12 of the 28 schools involved in the project from 2003–2005 hired new principals during that time period. Additionally, comparative analyses found differences between teachers and administrators in their perspectives on systemic work in their sites. Administrators were generally more optimistic about the development of important practices in their sites than were teachers. These distinctions between teachers and administrators suggested that progress through the model was uneven within sites and was more visible to administrators than to teachers.

Exploratory analyses of selected sites illuminated some of these findings. As seen in examinations of data for both Wrightsville, Oklahoma, and Roydale, Arkansas, differences between *Working Systemically* survey and interview results may have been due to differences in perspectives and knowledge of the work between school and district leadership and faculty. In the case of Wrightsville, leaders appeared to lose some of their focus on data-driven decision making and began to attend to other issues. However, these changes appeared not to affect negatively the regular work of the site in that the perceptions of teachers and other staff, as measured by the survey, continued to be positive.

In Roydale, however, the lack of cohesion between the school and district leadership working with field staff and the larger school and district staffs caused positive developments on the part of the leadership to have little effect on teachers' perceptions and assessments of the progress of the district. In both cases, divergence between the work of leaders in the district and that of faculty and staff led to different outcomes. In Wrightsville, this turned out to be positive because most staff in the district were shielded from leaders' loss of focus. However, staff in Roydale appear to have been somewhat unaware of positive developments emanating from district leadership. Such differences can obviously have substantial effects on the outcomes of districts both in working systemically and in improving student achievement.

In all, sites improved in working systemically, but those improvements were uneven and relatively limited. At the same time, there are encouraging findings. Sites exhibited growth in setting clear expectations and priorities on improving student achievement, focusing professional development on student learning needs, and aligning curriculum to state standards. The following section examines the extent to which these improvements resulted in improved student outcomes.

V. STUDENT ACHIEVEMENT OUTCOMES

Did student achievement increase as districts and schools built their capacity to work systemically?

Results from previous sections indicated that while staff utilized similar strategies to develop systemic work, there was variation in how these strategies were implemented owing to different site needs and characteristics. In addition, sites varied in their development of systemic work. Correlational analysis of *Working Systemically* survey results and interview scores suggested a strong positive relationship between the two measures of working systemically, thus providing some corroborative evidence for educators' perceptions of systemic work in a site.

This section investigates the relationship between working systemically and student achievement. The *Working Systemically* model proposed to boost achievement by increasing school and district support of curriculum, instruction, and assessment. Positive trends in achievement concurrent with increases in systemic work would, therefore, provide support for the work. It is important to note, however, that a correlation between improved achievement and development of systemic work does not imply a causal link.

The data collected for analysis were the proportion of students in tested grades scoring at a specific performance level defined by state education standards.³⁰ For example, the Texas Education Agency publicly releases student test information as the number of students who meet minimum standards. Other states have more specific performance categories (e.g., advanced, basic, unsatisfactory). However, all of these states have also devised a "cut score" that specifies the minimum performance level students must achieve in order to meet state criteria for proficiency. Student results in this report are conveyed as the percentage of students in tested grade levels who have met the proficiency standard set by the state (see Section II).

³⁰ Ideally, student achievement analyses would focus on test scores reported at the individual level. However, a strict interpretation of the Family Educational Rights and Privacy Act (FERPA) by several states has restricted access to this level of data. Consequently, analyses in this report utilize non-parametric statistics (described in Section II) that are more appropriate for the type of categorical information currently available from state departments of education.

Analytic results presented in this section help us answer the following three fundamental questions related to the overall research questions:

- Were there significant changes in the proportion of students in SEDL schools meeting state achievement standards over time?
- Was the proportion of students meeting state performance standards in SEDL schools significantly different than the proportion of students meeting state standards in similar schools over time?
- Was there a relationship between student achievement and systemic work as described by the *Working Systemically* survey and site interview data?

Results are organized around these three sub-questions. The first section addresses sub-questions 1 and 2, providing an overview of results followed by a more detailed description of significant findings. The data reflect the focus of SEDL's work in a particular site. For example, if a school concentrated their work with SEDL on reading, then only student achievement data related to reading, English/language arts, or literacy were used in the analysis. This approach more accurately reflects the areas addressed by the *Working Systemically* model at a particular school or site.

The second section examines findings for sub-question 3. Analyses for this question incorporated a cross-site analysis of achievement data, *Working Systemically* survey data, and interview scores.

Supplementary Analyses

In addition to binary analyses (met standard vs. not met standard) conducted across all states, analyses were also carried out for SEDL schools in three states reporting more than two performance categories (Arkansas, Louisiana, and Oklahoma). The purpose of these analyses was to determine whether there were significant changes in the proportion of students meeting specific performance levels over time. The rationale for this approach has to do with the possibility that while there might not be significant changes in the overall "met standard" category, more subtle longitudinal shifts might be evident in particular categories like "basic or

“advanced.” For example, combining three levels, like advanced, proficient, and basic, into a “met standard” category potentially masks substantive changes in the “basic” category from one year to the next.

Chi-square analyses were used to examine whether there were overall longitudinal differences across all performance categories for each tested grade level in SEDL sites. In cases where the overall chi-square statistic indicated a significant change, multiple pairwise comparisons were conducted for each performance category for years included in the analysis.³¹ For instance, in Arkansas, researchers compared student proportions in Forked River High School B for the basic category (2003 vs. 2004), advanced category (2003 vs. 2004), and so on.

Results of these analyses did not indicate any consistent patterns in the performance categories across SEDL schools over time. Indeed, out of 111 possible pairwise contrasts spanning 2003 through 2005 for Arkansas, Louisiana, and Oklahoma, no contrasts were statistically significant. In light of these results, further discussion of achievement in this section focuses on the proportion of students meeting and not meeting state standards because this was and continues to be the most relevant determinant of whether a school meets AYP requirements.

Findings

Student Achievement Results

Student achievement was mixed across sites and schools. Thirteen grade levels increased the proportion of students meeting the state standards in either reading or math; eight decreased. One grade showed no change at all.

³¹ See Glass and Hopkins (1996) for a description of multiple comparisons using pairwise contrasts for proportions.

Table V.1 provides a summary of achievement results for all sites included in this analysis.

Table V.1 Summary Table of Achievement Across Sites

State	School	Met Standard		
		2003	2004	2005
Arkansas	1. Grisham–Elementary School C, Grade 4 Math	34%	37%	–
	2. Grisham–Middle School A, Grade 6 Math	5%	10%	–
	3. Forked River–Elementary School B, Grade 4 Literacy	43%	42%	–
	4. Forked River–Elementary School B, Grade 6 Literacy	34%	39%	–
	5. Forked River–High School B, Grade 8 Literacy	24%	39%	–
Louisiana	6. Athens–Middle School A, Grade 8 Math	72%	78%	66%
	7. Bayou City–Middle School B, Grade 8 Math	71%	71%	59%
	8. Bayou City–High School A, Grade 10 Math	72%	75%	71%
	9. Highway Junction–Elementary School A, Grade 4 English/Language Arts	58%	53%	55%
	10. Highway Junction–Elementary School B, Grade 4 English/Language Arts	71%	71%	63%
New Mexico	11. Desert Hills–Elementary School A, Grade 4 Reading	35%	40%	–
	12. Desert Hills–Middle School A, Grade 8 Reading	43%	48%	–
	13. Farmville–Elementary School B, Grade 4 Reading	23%	28%	–
	14. Farmville–Elementary School C, Grade 4 Reading	30%	39%	–
	15. Farmville–Middle School C, Grade 8 Reading	32%	38%	–
Oklahoma	16. Bricktown–Middle School A, Grade 8 Reading	72%	77%	–
	17. Brookesville–Elementary School D, Grade 4 Reading	39%	39%	–
	18. Brookesville–Middle School B, Grade 8 Reading	62%	57%	–
	19. Wrightsville–Elementary School B, Grade 4 Reading	53%	88%	–
Texas	20. Pineland–Elementary School A, Grade 3 Reading	84%	87%	92%
	21. Pineland–Elementary School A, Grade 4 Reading	80%	78%	68%
	22. Pineland–Elementary School A, Grade 5 Reading	79%	84%	84%

Table V.2 highlights results from the previous table that indicated significant differences in the proportion of students meeting the state standards from year to year.

Table V.2 Summary Table of Significant Changes Over Time

School	<i>Met Standard*</i>		
	2003	2004	2005
2. Grisham–Middle School A, Grade 6 Math, AR	5%	10%	–
6. Athens–Middle School A, Grade 8 Math, LA	72%	78%	66%
7. Bayou City–Middle School B, Grade 8 Math, LA	71%	71%	59%
19. Wrightsville–Elementary School B, Grade 4 Reading, OK	53%	88%	–
20. Pineland–Elementary School A, Grade 3 Reading, TX	84%	87%	92%
21. Pineland–Elementary School A, Grade 4 Reading, TX	80%	78%	68%

* $p \leq .05$

The bold font illustrates the pair or pairs of proportions that were significantly different for each grade-level analysis. For example, the five-percentage-point increase for Grisham–Middle School A met the criteria for statistical significance.

Table V.3 presents data on significant differences between SEDL school results and comparison school results.

Table V.3 Summary Table of Significant Differences Between SEDL School Results and Comparison Group Results

School	<i>Met Standard*</i>		
	2003	2004	2005
5. Forked River–High School B, Grade 8 Literacy, AR	24%	39%	–
7. Bayou City–Middle School B, Grade 8 Math, LA	71%	71%	59%
19. Wrightsville–Elementary School B, Grade 4 Reading, OK	53%	88%	–

* $p \leq .05$

In this case, the bold font points out instances where the proportion of students meeting state-defined benchmarks for the SEDL school was significantly different than the corresponding proportion among matched comparison schools during the same year.

Across 22 grade-level analyses, six cases suggested statistically significant differences among results. Three grade levels (numbered 2, 19, and 20 in Table V.2) had statistically

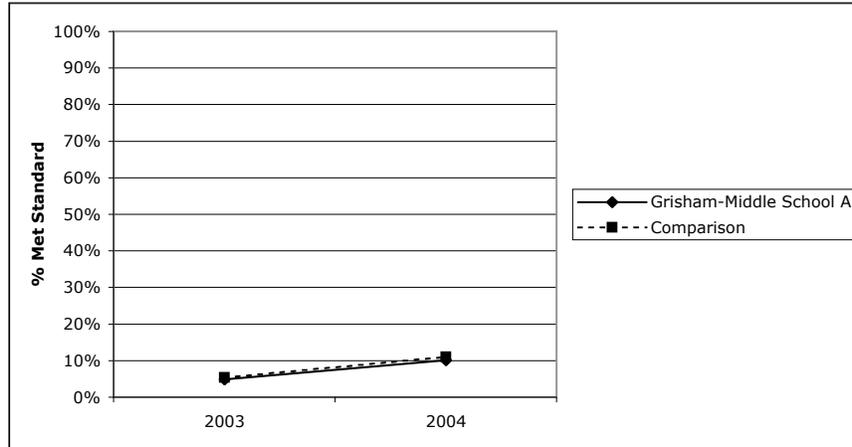
significant increases while three others (6, 7, and 21) demonstrated significant decreases. Only one grade (19) showed both a significant increase in the proportion of students meeting proficiency standards *and* a significant difference between SEDL school performance and the performance of the comparison schools (the latter is reported in Table V.3). Statistical analyses, including contingency tables, for all grade levels are included in Appendix F.

Researchers were unable to find suitable matches (primarily regarding initial achievement) for Forked River–High School B and Bayou City–Middle School B. Significant differences between these schools and their respective comparison groups were likely an artifact of the sampling strategy. Although this still permitted us to examine achievement trends for Bayou City–Middle School B, longitudinal results for Forked River–High School B did not support further examination because results did not indicate significant change.

Significant increases in the proportion of students meeting proficiency standards were spread across three states (Arkansas, Oklahoma, and Texas). The Arkansas case reflected achievement in math while the other two focused on reading. Achievement trends for these schools are presented below.

Figure V1 illustrates two groups (Grisham–Middle School A vs. comparison group) with extremely low aggregate achievement levels. Despite these low achievement levels, both groups had similar increases in the proportion of students meeting proficiency standards from 2003 to 2004.

Figure V.1 Grade 6 Math, AR, Grisham–Middle School A



Although both lines in the graph indicated positive changes, only the SEDL school increase was significant at $\alpha = .05$. While both groups have nearly identical percentage gains, the actual numerical increase was smaller for the comparison group ($N = 4$) relative to the number tested in that site than for the SEDL school ($N = 13$). The chi-square procedure is sensitive to sample size, which helps explain why the SEDL school meets the criteria for significance while the composite school does not.³² Despite the rise in achievement, student performance remained critically low in this site, highlighting the challenge facing many of these schools.

³² As total sample size decreases, the number of cases needed to impact proportions also decreases. For instance, a shift of two persons into the “met standard” category in a group of four evenly split between “met standard” and “did not meet standard” leads to a 100% passing rate. While this is a 50 percentage point increase, the small sample size precludes statistically significant findings ($\chi^2 = 2.67$). However, if 15 people shifted to met standard in a group of 30 evenly split between meeting standard and not meeting standard, the resulting 50 percentage point increase would be statistically significant ($\chi^2 = 20.00$). This explains how identical percentage increases might have different results in terms of statistical significance.

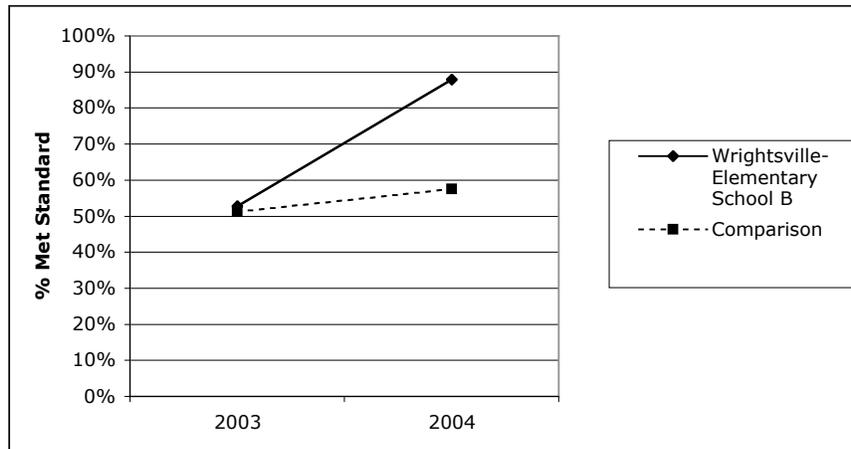
Table V.4 Chi-Square Analysis, Grade 6 Math, AR, Grisham–Middle School A

Comparison of Proportion of Students Meeting Standard	Sig.
SEDL Change (2003 to 2004)	*
Comparison Group Change (2003 to 2004)	ns
2003 SEDL vs. 2003 Comparison Group	ns
2004 SEDL vs. 2004 Comparison Group	ns

* $p \leq .05$

Wrightsville–Elementary School B made significant gains in the percentage of students meeting the Oklahoma state standards. This gain was in contrast to the comparison group, which saw a much smaller increase over the same period.

Figure V.2 Grade 5 Reading, OK, Wrightsville–Elementary School B



Chi-square statistics reveal that the difference between the comparison group and the SEDL school was also significant.

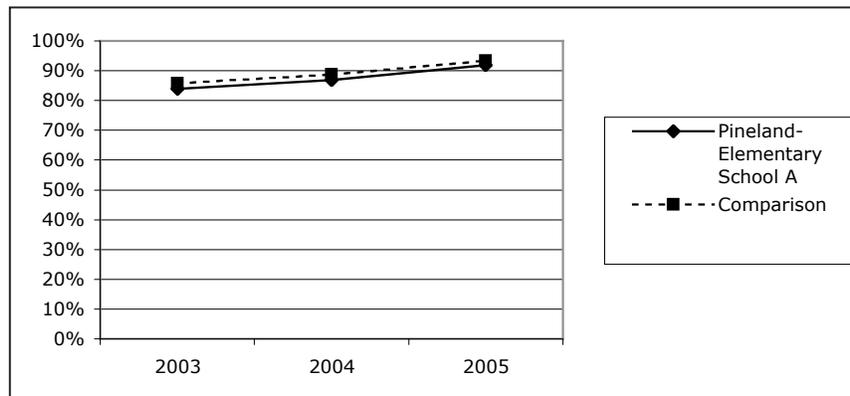
Table V.5 Chi-Square Analysis, Grade 5 Reading, OK, Wrightsville–Elementary School B

Comparison of Proportion of Students Meeting Standard	Sig.
SEDL Change (2003 to 2004)	*
Comparison Group Change (2003 to 2004)	ns
2003 SEDL vs. 2003 Comparison Group	ns
2004 SEDL vs. 2003 Comparison Group	*

* $p \leq .05$

Figure V.3 compares grade 3 reading results for Pineland–Elementary School A and the matched comparison group. The graph depicts two parallel, nearly identical lines showing a gradual increase over 3 years.

Figure V.3 Grade 3 Reading, TX, Pineland–Elementary School A



Although there was a percentage increase for both groups over time, only the SEDL school registered a significant change from 2003 to 2005. Supporting the visual evidence, statistical analyses did not indicate any significant differences between the two groups over time. Results for the chi-square analyses are presented in Table V.6.

Table V.6 Chi-Square Analysis, Grade 3 Reading, TX, Pineland–Elementary School A

Comparison of Proportion of Students Meeting Standard	Sig.
SEDL Change (2003 to 2004)	ns
SEDL Change (2004 to 2005)	ns
SEDL Change (2003 to 2005)	*
Comparison Group Change (2003 to 2004)	ns
Comparison Group Change (2004 to 2005)	ns
Comparison Group Change (2003 to 2005)	ns
2003 SEDL vs. 2003 Comparison Group	ns
2004 SEDL vs. 2004 Comparison Group	ns
2005 SEDL vs. 2005 Comparison Group	ns

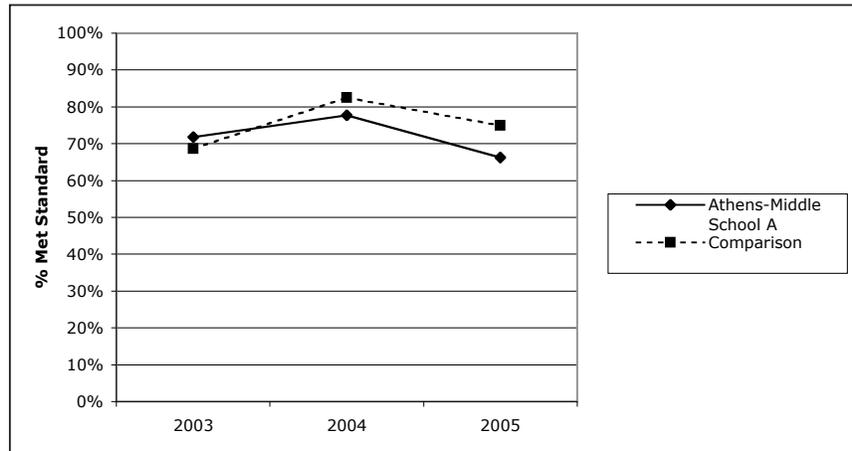
* $p \leq .05$

Among the three cases demonstrating increases over time, two sets of increases (Grisham–Middle School A and Pineland–Elementary School A) were not found to be significantly different over time when compared to achievement patterns for similar schools.

Statistically significant decreases in achievement levels were found in three cases located in Louisiana and Texas. Two of the decreases occurred in math sites while the third decrease happened in reading.

Figure V.4 illustrates achievement trends for Athens–Middle School A and its matched comparison group. The comparison group experienced a significant increase in aggregate achievement from 2003 to 2004. A non-significant decline occurred from 2004 to 2005.

Figure V.4 Grade 8 Math, LA, Athens–Middle School A



Athens–Middle School A followed a similar pattern of achievement to the matched set. However, the SEDL school posted a statistically significant decline in students meeting state proficiency standards from 2004 to 2005. The gain in proficiency from 2003 to 2004 was not significant. In addition, the overall change from 2003 to 2005 for both groups was not significant.

Table V.7 Chi-Square Analysis, Grade 8 Math, LA, Athens–Middle School A

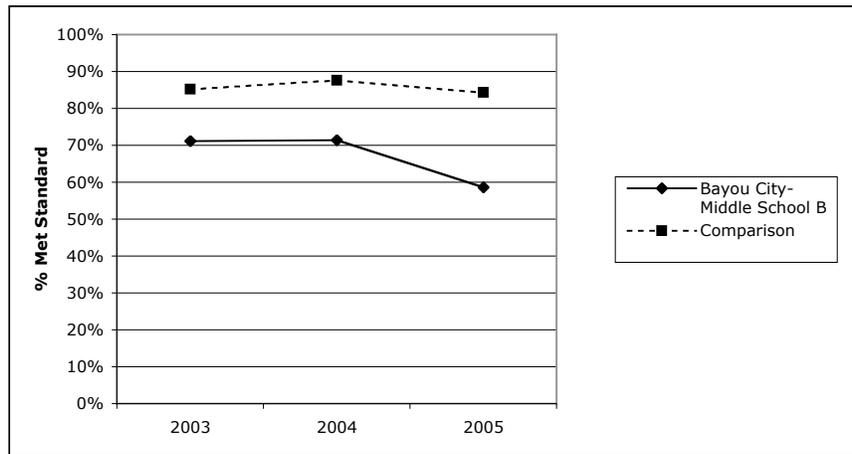
Comparison of Proportion of Students Meeting Standard	Sig.
SEDL Change (2003 to 2004)	ns
SEDL Change (2004 to 2005)	*
SEDL Change (2003 to 2005)	ns
Comparison Group Change (2003 to 2004)	*
Comparison Group Change (2004 to 2005)	ns
Comparison Group Change (2003 to 2005)	ns
2003 SEDL vs. 2003 Comparison Group	ns
2004 SEDL vs. 2004 Comparison Group	ns
2005 SEDL vs. 2005 Comparison Group	ns

* $p \leq .05$

The pattern of change for Athens–Middle School A was consistent with trends in the comparison group. However, while the difference in percent meeting the standard was significant from 2004 to 2005, the difference in proportions between the SEDL school and comparison group was not significant in any year.

Differences in achievement levels between the comparison group and SEDL school are apparent in Figure V.5. It is important to note that the researchers in this case were unable to find schools that matched the SEDL school on the initial achievement variables. Therefore, the differences shown in the graph were not surprising. More telling are the rates of change when they occurred.

Figure V.5 Grade 8 Math, LA, Bayou City–Middle School B



Through 2004, the rate of change between these groups was roughly parallel. However, between 2004 and 2005, Bayou City–Middle School B posted a significant decrease in the proportion of students meeting state proficiency standards. The comparison group remained relatively unchanged over the same period. Overall change was also statistically significant for the SEDL school from 2003 to 2005 but not for the comparison group.

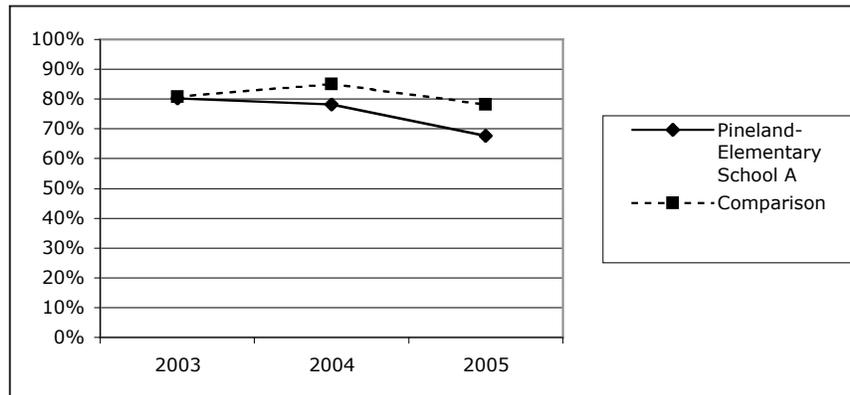
Table V.8 Chi-Square Analysis, Grade 8 Math, LA, Bayou City–Middle School B

Comparison of Proportion of Students Meeting Standard	Sig.
SEDL Change (2003 to 2004)	ns
SEDL Change (2004 to 2005)	*
SEDL Change (2003 to 2005)	*
Comparison Group Change (2003 to 2004)	ns
Comparison Group Change (2004 to 2005)	ns
Comparison Group Change (2003 to 2005)	ns
2003 SEDL vs. 2003 Comparison Group	*
2004 SEDL vs. 2004 Comparison Group	*
2005 SEDL vs. 2005 Comparison Group	*

* $p \leq .05$

Figure V.6 illustrates grade 4 achievement patterns for Pineland–Elementary School A and its matched comparison group.

Figure V.6 Grade 4 Reading, TX, Pineland–Elementary School A



Both groups demonstrated an overall decline in the proportion of students meeting the state standards. However, while changes within the comparison group were not statistically significant, the drop from 2003 to 2005 for Pineland–Elementary School was significant over time. Differences between the schools in each year were not statistically significant (see Table V.9).

Table V.9 Chi-Square Analysis, Grade 4 Reading, TX, Pineland–Elementary School A

Comparison of Proportion of Students Meeting Standard	Sig.
SEDL Change (2003 to 2004)	ns
SEDL Change (2004 to 2005)	ns
SEDL Change (2003 to 2005)	*
Comparison Group Change (2003 to 2004)	ns
Comparison Group Change (2004 to 2005)	ns
Comparison Group Change (2003 to 2005)	ns
2003 SEDL vs. 2003 Comparison Group	ns
2004 SEDL vs. 2004 Comparison Group	ns
2005 SEDL vs. 2005 Comparison Group	ns

* $p \leq .05$

While the previous three cases were characterized by significant declines, it is unclear whether these declines were endemic to the schools. In two cases, results for the comparison groups were not significantly different from achievement patterns for the SEDL schools. Although there was a statistically significant difference between Bayou City–Middle School B and its matched comparison group, this difference was at least partially attributable to sampling bias as suggested earlier.

Analyses of statistically significant grade-level achievement patterns emphasize the mixed nature of the overall results. For example, while Wrightsville–Elementary School B was the only school to demonstrate significant achievement increases that were also significantly different from a matched comparison group, interview and survey results for the site are not wholly consistent with this finding. Although the pattern of interview scores suggested an increase in working systemically, survey results demonstrated a significant decline.

Interview and survey results for the other five cases highlighted in this section offer similar patterns. For instance, as a site, Athens posts a statistically significant increase in perceived working systemically via the interview data (see Table IV.1). However, this stands in contrast to the significant decline in achievement in Athens–Middle School A portrayed in Figure V.4.

Case-by-case analysis of achievement and working systemically paints a murky picture of the relationship between these two sets of outcomes. So far, we have only been concerned with the six schools that exhibited statistically significant changes over time. However, sub-question 3 suggests a wider analysis, examining achievement and working systemically outcomes for all available sites and grade-level data.

Student Achievement and Working Systemically

The comparative grade-level achievement analyses revealed one significant change in the proportion of students meeting the state standards in reading or math. However, the question remains regarding the relationship, if any, between working systemically and student achievement. While data limitations precluded more complex analyses, it was possible to devise a blunt measure of this relationship.

To examine the correlation between working systemically and student achievement, the research team used data for Spring and Fall 2003 and 2004 because the research team had two sets of overlapping data during this period: 1) Spring 2003 achievement data and Fall 2003 interviews and 2) Spring 2004 surveys, Spring 2004 achievement results, and Fall 2004 interviews. The researchers relied on fall interviews because they were conducted early in the school year when SEDL's work was just beginning and respondents tended to reflect on the previous year's efforts.

Table V.10 presents the correlation between Spring 2003 achievement data and Fall 2003 interview results.³³ A one-tailed significance test was employed because the theoretical relationship between these measures supports a positive directional hypothesis.

³³ It is important to recall that n represents the total number of grade-level analyses. Site-level mean interview scores are assigned to each grade level in a particular site, consistent with the analytic strategy discussed in the methods section.

Table V.10 Correlation Between Interviews and Achievement, 2003

	Interviews
Student Achievement	.59*
<i>N</i>	22

* $p \leq .05$

The correlation analysis indicated a positive relationship between working systemically (as captured by interviews) and the proportion of students meeting proficiency standards.

Table V. 11 illustrates the correlations between two measures of working systemically and achievement for 2004.

Table V.11 Correlation Between Working Systemically Measures and Achievement, 2004

	Interviews	Survey
Student Achievement	.37*	.33
<i>N</i>	22	22

* $p \leq .05$

The correlation results indicated a continued significant relationship between interview results and achievement. *Working Systemically* survey results and achievement were not significantly correlated. Although the correlation between interview results and student achievement is weaker in 2004 versus 2003, a statistical test suggested that the difference between the two correlation coefficients was not statistically significant.

To supplement the overall correlation between interview results and student achievement data, researchers did further analyses between achievement and the subscales for the interviews (i.e., Expectations and Priorities, Alignment, Collaboration, and Professional Development). The purpose of this exercise was to determine whether particular working systemically activities were more closely related to achievement outcomes. Table V.12 and Table V.13 provide results for 2003 and 2004 data, respectively.

Table V.12 Correlations Between Achievement and Interview Subscales, 2003

Interview Sub-Scale	2003 Achievement
Expectations and Priorities	.26
Alignment	.49 *
Collaboration	.57 *
Professional Development	.55 *
<i>N</i>	22

* $p \leq .05$

Table V.13 Correlations Between Achievement and Interview Subscales, 2004

Interview Sub-Scale	2004 Achievement
Expectations and Priorities	.13
Alignment	.43 *
Collaboration	.15
Professional Development	.04
<i>N</i>	22

* $p \leq .05$

Although three of four correlations were significantly correlated with achievement in 2003, this pattern does not continue in 2004. Alignment scores, however, were significantly correlated with achievement in both years.

The correlation results for 2003 and 2004 provide some support for the existence of a relationship between development of systemic work and student achievement. Among specific activities, alignment had the closest relationship to achievement in both years included in the analyses. It is important to remember that correlation does not imply causality. Underlying any correlational analysis is the possibility of a spurious relationship. As discussed earlier in this section, the analyses provide a crude descriptive measure of a complex program. However, although limitations are certainly evident, the direction and significance of the correlations do fit the basic theoretical framework for this work.

Discussion

The static nature of the achievement trends could be attributed to any number of factors. A common issue raised regarding state accountability tests is the fact that they are generally administered only once a year (Stecher et al., 2003). Therefore, when examining longitudinal change, aggregate data from 2 consecutive years actually represent two different populations of students. One implication is that changes in outcomes (positive or negative) may simply reflect a change in the characteristics of the sample.

Another potential factor is school characteristics. The majority of the sites involved with SEDL were selected precisely because they were low-achieving schools in high-poverty locations. Two major issues in such districts include turnover (teacher and/or administrators) and student mobility. High levels of turnover, mobility, or both could have a substantial negative impact on achievement as well as undermine the beneficent influence of external efforts to build capacity (Mujis et al., 2004). Indeed, one of the schools in Louisiana that posted a significant decrease in achievement during 2004–2005 began the school year with a new principal.

Ceiling effects may have played a role in the Texas school. As demonstrated by Table V.1, Pineland–Elementary School A had relatively high proportions of students meeting state standards in all 3 years. Consequently, significant increases might be harder to detect.

Another possibility is that the distal nature of the intervention precludes the appearance of significant changes in student outcomes within the given timeframe. In their review of a Chicago reform effort emphasizing systemwide decentralization to facilitate school improvement, Bryk, Kerbow, and Rollow (1997) present a finding that lends credence to this scenario. After 5 years of implementation, a grade- and subject-level analysis of student achievement patterns found that the most prevalent trend was one of “no change” (Bryk et al., 1997).

As discussed in the limitations section, the use of non-parametric statistics in this report is a function of the type of data available. However, non-parametric statistics are not as powerful as parametric statistics for analyzing differences between groups. For this reason, it is important to recall the concept of practical significance. In other words, were there any apparent trends irrespective of statistical significance? Although there were only three increases that were statistically significant overall, 13 grade-level analyses showed increases. Results thus indicate that a majority of schools were “trending up.”

Of critical importance in this discussion is the fact that AYP focuses on the percentage of students meeting specific standards set by the states. One implication of this policy is that even small, non-significant percentage changes may have substantial implications for a school's AYP status. For many of the schools working with SEDL, any increase is a positive step.

A more general investigation of available data presented in the previous section suggests at least a partial relationship between interview results and achievement data across SEDL sites. In 2 consecutive years, interviews and achievement data were positively correlated.

Unlike the survey, which concentrated on more global perceptions and attitudes regarding systemic work, the interview analyses were focused on educator practice. Higher ratings on particular sub-scales derived from more concrete examples of activities and behaviors that supported systemic work. Theoretically, these actions and behaviors are more directly associated with what is happening to support classroom practice and learning. Therefore, it makes sense that in both years, the interview results would have a stronger relationship to achievement than survey results.

Clearly, correlation analyses are not a definitive answer to whether there is a link between working systemically and student outcomes. Considering the complexity and variation of the *Working Systemically* model and the coarseness of the achievement data, these correlation results are encouraging and support further investigation into the nature of the relationship.

VI. CONCLUSION

The prevalence nationwide of test-based accountability systems provides the context for SEDL's work. The goal of these accountability systems is to improve student achievement by focusing instruction on specific state academic standards. While there are numerous programs that attempt to guide and inform this process at the school and classroom level, SEDL relied on a systemic approach, addressing all levels and components of districts in support of better instruction and improved student outcomes.

Findings

The research component of SEDL's work examined three questions related to the *Working Systemically* model. These questions related to the strategies field staff used to build educators' capacity to work systemically, the degree to which schools and districts increased their capacity to work systemically, and whether there was a relationship between increased capacity and student achievement. Analyses resulted in three general findings: (1) field staff employed three core strategies to implement plans that would build capacity for systemic work; (2) there was limited growth in working systemically although context was an important factor supporting sites' progress; and (3) longitudinal achievement patterns were mixed for working systemically sites. However, there was a small but significant relationship between working systemically measures and student achievement.

Field Staff Strategies Focused on Instructional Leadership

Field staff used a fairly coherent set of strategies to develop systemic work in the sites. This report examined those designed to have the most direct impact on the classroom. Three core strategies were central: (1) finalizing action plans and ensuring all initiatives in the site were coherent with its goals; (2) continuing to build educators' capacities, especially through meeting structures, coaching, and via the PTLC; and (3) helping leaders monitor activities related to the PTLC and progress on related goals at the district level. Use of the PTLC within and across

grade levels strengthened participants' knowledge and skills and catalyzed cross-level communication and collaboration through the ways field staff encouraged leaders to monitor it. Despite the emphasis these three core strategies placed on classroom-level practice, all of them stressed the importance of instructional leadership. The accountability of educators to the levels above their own proved crucial to conducting the work supported by the field staff.

Limited Growth in Systemic Work

Analyses indicated that sites did make progress in working systemically. They developed skills and practices that enabled them to integrate and direct the various facets of their school systems toward achieving student learning goals. The amount of growth was generally small, however, and there was variation in outcomes among sites and among groups of respondents. Teachers and administrators diverged on many measures, including the extent to which student learning was made a priority and the quality of professional development in their districts.

Differences in patterns of participation in SEDL's work and knowledge of new initiatives resulting from the work most likely contributed to these differences. One of the primary challenges encountered in the project had to do with the fact that this work was conceived as a research *and* development project. In other words, the *Working Systemically* model was refined by field staff as they gained knowledge and insight about what was working in the sites. This tension was manifest in attempts by field staff to balance the demands of a dynamic systemic reform model and the needs of a particular school district. In a practical sense, variability was built into the process.

The implication of the development component of the process was its impact on consistency of field staff practice. Strategies described in site contact records reveal shifting priorities across sites. This variation was inevitable due to the context-specific nature of the model. While the broad strategies used by field staff were similar, utilization of strategies and

activities differed according to site needs. These differences in the implementation of the model across sites underscore the variation expressed in working systemically measures.

Another factor contributing to differences was the relative competency and skill level across sites. For example, the model relied heavily on teams to set priorities and make data-driven decisions about resource utilization toward improving student outcomes. However, teacher and administrators first had to learn how to collaborate before they could focus their attention on creating coherence across the system. Even upon exiting the sites, field staff reported some educators still lacked basic knowledge of how to evaluate and utilize research, technical assistance, and other resources.

Mixed Student Achievement Patterns

Limited gains in working systemically may be reflected in student outcomes. Overall, grade-level achievement results did not indicate a pattern of significant change among schools participating in the project. Thus, the initiative did not provide compelling evidence of impacts on long-term criterion-referenced indicators of academic achievement. However, significant correlations between working systemically measures and academic achievement, particularly for alignment, provide non-definitive but encouraging evidence that progress on these short-term outcomes may be important in terms of academic achievement.

Recommendations for Systemic Reform

The question of whether systemic work is linked to increased student achievement highlights the multiple demands intrinsic to systemic reform, particularly in low-performing districts. Fundamentally, districts must know how to recognize critical needs (e.g., improving achievement) and be able to devise and implement plans to meet those needs. Honig and Hatch (2004) envision this process as an informed negotiation between school leaders and district administrators while balancing school goals and policy demands. A tacit assumption of this process, however, is that all stakeholders have the capacity to be fully engaged.

Both field staff and educators in the sites reported that they believed more time to develop systemic work would be necessary in order to affect student achievement. Other school improvement efforts have found that at least 5 years are needed to establish initiatives into school practice and realize gains in student achievement (Berends et al., 2005; Bryk et al., 1997). Questions remain, though, about the reasons underlying the need for more time. Our analyses suggest that one reason may be that the implementation of improvement initiatives requires knowledge, skills, and resources that schools and districts do not always have. Hatch (2001) underscores this paradox facing low-performing schools:

While they [schools] may need help to develop capacity to raise standards and change many aspects of their operations, the implementation of improvement programs is difficult precisely because schools lack the capacity to change. (p. 47)

In putting the *Working Systemically* model into practice, field staff focused on systemic processes such as collecting data, setting goals and objectives, and planning initiatives. It became apparent, however, that teachers and administrators lacked skills and resources to use these processes effectively. Field staff then began to emphasize coaching and modeling in their work to build educators' capacity while implementing new programs and initiatives being developed through districts' work with the *Working Systemically* model. For example, when working with both teachers and administrators, collaborative planning and decision making was a common strategy used to set goals and create coherence in the sites. Many educators, however, had had little experience in working collaboratively and needed training on how to use their planning time productively to achieve explicit goals.

Issues such as materials, staffing, scheduling, and other resources also need to be addressed. Many sites lacked instructional materials that were aligned with their states' standards. Some school principals also found it challenging to incorporate classroom walkthroughs into their schedules alongside their operational and managerial responsibilities. These principals often did not have assistant principals in their buildings to whom they could delegate administrative duties. Some sites faced challenges in sustaining leadership teams or

conducting professional development activities when these meetings could not be scheduled during the regular school day.

The process of systemic reform then becomes two-fold. First, reform efforts have to focus on increasing school and district capacity. Second, program personnel have to help educators utilize capacity coherently and effectively at various system levels. In particular, improvement initiatives need to focus explicitly on improving skills related to classroom instruction and developing school and district infrastructures to support and sustain coherence in instructional programs (Newmann et al., 2001). This is evident in the findings about the relationship between alignment of curriculum to state standards and student achievement.

Alignment was positively and consistently related to improved achievement. Other areas of practice (Expectations and Priorities, Collaboration, and Professional Development) were not as consistently or strongly associated with student achievement. This does not suggest that improvement efforts should focus solely on aligning curriculum to standards. Rather, it indicates that directly addressing curriculum and instruction may lead to improved student learning because alignment was the area of practice most directly related conceptually to teaching and learning.

As the *Working Systemically* model suggests, improvements to curriculum and instruction do not occur in isolation from other school and district practices. Other areas of the system must support such improvements and should be directed toward improved teaching and learning. Staffing practices, for example, must focus on recruiting and retaining teachers who have the content and pedagogical knowledge to implement instructional strategies that meet students' needs. So too do administrators need to be able to evaluate the quality of instruction and support teachers in improvement efforts.

Recommendations for Research

The recommendations for research that follow from SEDL's experience with this research and development project relate both to content and methodological concerns. The recommendations for practice, grounded in the research findings, direct attention to topics for future study. Practitioners have yet to learn from researchers how best to address the capacity issues that emerged in this work. Methodological issues relate to capacity levels as well. Any systemic or school improvement project is needed primarily in sites characterized by low performance. Doing research generally, but especially in these sites, means educating participants about the research procedures as part of an ongoing recruitment process that promotes commitment and therefore fidelity to the intervention.

Building Capacity for Improvement

Systemic reform is by its nature multilevel and thus calls for a dual focus in the kinds of technical assistance provided. This research found that progress in implementing some of the work outlined in district action plans was hampered by a lack of resources: written curriculums, appropriately focused benchmark assessments, and professional development consistent with careful considerations of which areas of capacity growth should be priorities. The field staff certainly facilitated efforts in this area. Some teachers and school-level administrators, however, reported feeling peripheral to SEDL's work until it more directly engaged issues of instruction. Researchers need to examine more systematically what kinds of technical assistance will support administrators' leadership capacity as it simultaneously improves instruction.

The PTLC was theorized to boost student achievement by promoting alignment of instruction to a standards-based curriculum and assessment. Future research should elaborate on exactly why alignment might boost student achievement (Newmann et al., 2001). This work addressed some aspects of this question. The measures used in this research primarily addressed alignment of curriculum to state standards. Curriculum and state standards, as well as state-

mandated assessments, are written documents that can be physically compared. The “black box” in the alignment process is instruction.

Some existing work (Marzano, 2001) outlines instructional strategies that are likely to engage students and bring about the higher-order knowledge and skills increasingly called for in state standards. Future research needs to emphasize more formal observations of classroom practice that will systematically document which instructional strategies convey a standards-based curriculum in ways that will improve students' performance on aligned assessments. Researchers can also focus on how best to ensure that teachers know how to implement these strategies.

Methodological Considerations

Site selection is critical to school improvement initiatives. The majority of sites participating in this work were selected due to low performance on state standards-based achievement tests. However, the focus on low-performing schools introduced other issues particularly endemic to these environments including limited knowledge of what engagement in a research project entails, concurrent school improvement programs, and administrator turnover.

Knowledge of what participation in a research project entails is generally limited to scholars. To be evaluated scientifically, any school improvement intervention must therefore incorporate strong incentives for participants, leaders especially, to ensure continued commitment to it. This process of recruitment has to be ongoing in some sense, similar to the continuing emphasis on the rights of human subjects in any ethical social science research.

Projects sited in low-performing systems need to pay special attention to the external environment in which the work is conducted. Educators and field staff in this project experienced a negotiation among the demands of several school improvement programs typical of low-performing systems especially. Although SEDL had intentionally developed selection criteria to minimize the impact of external programs, demands on low-performing schools and districts

often led to the adoption of reform initiatives aimed at bolstering student achievement. These programs included Success for All, Reading First, and other local and state programs.

The role of leadership in directing and monitoring school improvement programs is crucial to their success (Mujis, 2004). Changes in leadership can greatly influence the direction of a reform process positively or negatively. In either case, a period of adjustment necessarily follows. Several sites involved with SEDL experienced these situations. In 16 sites, shifts occurred at the school level while four other sites had to adjust to new superintendents. Ideally, fostering an awareness among leaders of the role they play in the success of research interventions would be part of the coaching mentioned above.

In order for substantially low-performing districts to successfully implement systemic reform, resources must be directed toward helping educators at all levels attain a threshold of competency in needed areas. This strategy is consistent with observations that improved capacity is critical to meeting the accountability demands in underperforming schools and districts (Goertz, 2005). Implicit in this strategy is the need to address capacity for all stakeholders (e.g., teachers, school administrators, district administrators) concurrently. It is not enough to invest time and resources on one group over another. Long-term success is largely dictated by how well all the parts of the educational system work toward a common goal. Continued improvement is only possible when educators at all levels have the capacity to engage effectively in the reform process.

REFERENCES

- Arkansas Department of Education. (2004). *Arkansas comprehensive testing, assessment, and accountability program: Definitions*. Retrieved August 11, 2004, from http://arkedu.state.ar.us/actaap/definitions/definitions_p1.htm#Levels
- Arkansas Department of Education. (2005). *Arkansas consolidated school improvement planning (ACSIP) handbook*. Retrieved June 23, 2005, from http://arkedu.state.ar.us/whats_new/pdf/acsip_handbook_april05.pdf
- Berends, M., Bodilly, S., & Kirby, S. N. (2005). Reforming whole schools: Challenges and complexities. In J. Petrovich & A. S. Wells (Eds.), *Bringing equity back: Research for a new era in american educational policy* (pp. 161–194). New York: Teachers College Press.
- Borman, G. D., & Hewes, G. (2001). *The long term effects and cost-effectiveness of Success For All*. Baltimore, MD: Center for Research on the Education of Students Placed at Risk, Johns Hopkins University.
- Bryk, A. S., Kerbow, D., & Rollow, S. (1997). Chicago school reform. In D. Ravitch & J. P. Vitteritti (Eds.), *New schools for a new century: The redesign of urban education* (pp. 164–200). New Haven, CT: Yale University Press.
- Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic schools and the common good*. Cambridge, MA: Harvard University Press.
- Center for Mental Health in Schools. (2001). *Organization facilitators: A change agent for systemic school and community changes*. Los Angeles: Author.
- Cohen, D. K., & Ball, D. L. (1990). Relations between policy and practice: A commentary. *Educational Evaluation and Policy Analysis*, 12(3), 331–338.
- Cohen, D. K., & Ball, D. L. (1999). *Instruction, capacity, and improvement*. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.
- Corcoran, T., & Christman, J. B. (2002). *The limits and contradictions of systemic reform: The Philadelphia story*. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.
- Datnow, A., Borman, G., & Stringfield, S. (2000). School reform through a highly specified curriculum: Implementation and effects of core knowledge sequence. *The Elementary School Journal*, 101(2), 167–192.
- Edmonds, R. (1979). Effective schools for the urban poor. *Educational Leadership*, 37(1), 15–23.
- Elmore, R. F., Abelman, C. H., & Fuhrman, S. F. (1996). The new accountability in state education reform: From process to performance. In H. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 65–98). Washington, DC: Brookings Institution Press.
- Erickson, F., & Gutierrez, K. (2002). Culture, rigor, and science in educational research. *Educational Researcher*, 31(8), 21–24.

- Erzberger, C., & Kelle, V. (2003). Making inferences in mixed methods: The rules of integration. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 457–488). Thousand Oaks, CA: Sage Publications.
- Feuer, M. J., Towne, L., & Shavelson, R. J. (2002). Scientific culture and educational research. *Educational Researcher*, 31(8), 4–14.
- Fullan, M. (2001). *The new meaning of educational change*. New York: Teachers College Press.
- Glass, G., & Hopkins, J. D. (1996). *Statistical methods in education and psychology*. Boston: Allyn & Bacon.
- Goertz, M. E., & Massell, D. (2005). *Holding high hopes: How high schools respond to state accountability policies*. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.
- Greene, J. C., & Caracelli, V. J. (1997). *Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms*. San Francisco: Jossey-Bass.
- Hall, G. E., & Hord, S. M. (2001). *Implementing change: Patterns, principles and potholes*. Boston: Allyn and Bacon.
- Hallinger, P., & Murphy, J. (1986). The social context of the effective school. *American Journal of Education*, 94, 328–355.
- Hatch, T. C. (2001). It takes capacity to build capacity. *Education Week*, 20, 44, 47.
- Honig, M. L., & Hatch, T. C. (2004). Crafting coherence: How schools strategically manage multiple, external demands. *Educational Researcher*, 33(8), 16–30.
- Laczko-Kerr, I., & Berliner, D. C. (2002). The effectiveness of "Teach for America" and other under-certified teachers on student academic achievement: A case of harmful public policy, *Education Policy Analysis Archives* 10(37). Retrieved August 11, 2004 from <http://epaa.asu.edu/epaa/v10n37>
- LeCompte, M. D., & Schensul, J. J. (1999). *Analyzing and interpreting ethnographic data: Ethnographer's toolkit* (Vol. 5). Walnut Creek, CA: Altamira Press.
- Leithwood, K., & Menzies, T. (1998). A review of research concerning the implementation of site-based management. *School Effectiveness and School Improvement*, 9(3), 233–285.
- Louisiana Department of Education. The seven links of LINCS: A professional development process. In *Professional Development LINCS*. Retrieved August 19, 2005, from <http://www.doe.state.la.us/lde/pd/842.html>
- Louisiana Department of Education. (2004). *What is LEAP 21?* Retrieved August 11, 2004, from <http://www.doe.state.la.us/lde/uploads/1703.pdf>
- Marzano, R. J., Pickering, D. J., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for improving student achievement*. Alexandria, VA: Association for Supervision and Curriculum Development.

- Massell, D. (1998). *State strategies for building capacity in education: Addressing the needs of standards-based reform*. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.
- Mujis, D., Harris, A., Chapman, C., Stoll, L., & Russ, J. (2004). Improving schools in socioeconomically disadvantaged areas: A review of research evidence. *School Effectiveness and School Improvement*, 15(2), 149–175.
- New Mexico Department of Education. (2004). *General test specifications: New Mexico standards-based assessment, grades 4 and 8*. Retrieved September 28, 2004, from <http://www.ped.state.nm.us/div/ais/assess/dl/crt.descriptors.pdf>
- Newmann, F. M., Smith, B., Allensworth, E., & Bryk, A. S. (2001). Instructional program coherence: What it is and why it should guide school improvement policy. *Educational Evaluation and Policy Analysis*, 23(4), 297–321.
- Newmann, F. M., & Wehlage, G. G. (1995). *Successful school restructuring: A report to the public and educators*. Madison, WI: Center on Organization and Restructuring of Schools, University of Wisconsin.
- O'Day, J., Goertz, M. E., & Floden, R. E. (1995). *Building capacity for education reform*. New Brunswick, NJ: Consortium for Policy Research in Education, Rutgers University.
- Odden, A. (1998). *Creating school finance policies that facilitate new goals*. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.
- Oklahoma Department of Education. (2004). *Office of Accountability and Assessment performance level descriptors*. Retrieved August 11, 2004, from <http://www.sde.state.ok.us/home/defaultie.html>
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (Third ed.). Orlando, FL: Harcourt Brace.
- Puma, M., Raphael, J., Olson, K., & Hannaway, J. (2000). *Putting standards to the test: A design for evaluating the systemic reform of education*. Washington, D.C.: Urban Institute.
- Purkey, S. C., & Smith, M. S. (1983). Effective schools: A review. *The Elementary School Journal*, 83(4), 427–452.
- Sashkin, M., & Egermeier, J. (1993). *School change models and processes: A review and synthesis of research and practice*. Washington, D.C.: U.S. Department of Education.
- Sizer, T. R. (1996). *Horace's hope*. New York: Houghton Mifflin.
- Slavin, R. E., Madden, N. A., Dolan, L. J., Wasik, B. A., Ross, S., Smith, L., et al. (1996). Success for all: A summary of research. *Journal of Education for Students Placed at Risk*, 1(1), 41–76.
- Smith, M. S., & O'Day, J. (1990). Systemic school reform. *Politics of Education Association Yearbook*, 233–267.
- Southwest Educational Development Laboratory. (2000). *Creating knowledge to build high-performing learning communities: A proposal to serve as the regional educational laboratory for the southwestern region*. Austin, TX: Author.

- Southwest Educational Development Laboratory. (2002). *A description of SEDL's intensive sites: Working systemically to create high-performing learning communities*. Austin, TX: Author.
- Southwest Educational Development Laboratory. (2003). *The development and refinement of SEDL's Working Systemically model*. Austin, TX: Author.
- Southwest Educational Development Laboratory. (2004). *Alignment in SEDL's Working Systemically model*. Austin, TX: Author.
- Spillane, J. (2005). *Standards deviation: How schools misunderstand education policy*. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania.
- Spillane, J. P., & Thompson, C. L. (1997). Reconstructing conceptions of local capacity: The local education agency's capacity for ambitious instructional reform. *Educational Evaluation and Policy Analysis*, 19(2), 185–203.
- Stecher, B., Hamilton, L., & Gonzalez, G. (2003). *Working smarter to leave no child behind: Practical insights for school leaders*. Santa Monica, CA: RAND.
- Teddlie, C., & Tashakkori, A. (2003). Major issues and controversies in the use of mixed methods in social and behavioral sciences. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 3–50). Thousand Oaks, CA: Sage Publications.
- Texas Education Agency. (2004). *Interpreting assessment results: Spring 2004*. Retrieved August 23, 2005, from <http://www.tea.state.tx.us/student.assessment/resources/guides/interpretive/TAKS.pdf>
- Yin, R. K. (1984). *Case study research: Design and methods* (Vol. 5). Thousand Oaks, CA: Sage Publications.

**APPENDIX A:
SITE PROFILES**

Arkansas Development Sites

District Information

District	Type Location	Schools	Teacher FTEs	Students	Expenditure per Pupil	Focus of SEDL's Work
Forked River	Urban Fringe of Mid-size City	2	59	695	\$6,370	Reading

Source: Common Core of Data, National Center for Education Statistics, retrieved October 5, 2005, from <http://nces.ed.gov/globallocator>. Per pupil expenditure are for 2002–2003 school year; all other indices are from 2003–2004 school year.

School Information

<i>General School Information</i>					<i>Student Demographics</i>					
School	District	Grade Range	Teacher FTEs	Students	Free/Reduced Lunch	Hispanic	African American	Asian American	White	American/Alaskan Indian
Elementary B	Forked River	PK–6	31	375	71.5%	4%	30%	0%	67.1%	0%
High School B	Forked River	7–12	24	320	67.5%	2.2%	35.3%	0%	62.2%	0%

Source: Common Core of Data, National Center for Education Statistics, retrieved October 5, 2005, from <http://nces.ed.gov/globallocator>.

Louisiana Development Sites

District Information

District	Type Location	Schools	Teacher FTEs	Students	Expenditure per Pupil	Focus of SEDL's Work
Athens	Large Town	32	1,027	14,201	\$6,802	Math
Bayou City	Urban Fringe of Mid-Size City	29	1,152.9	14,872	\$7,091	Reading
Highway Junction	Small Town	7	159.1	2,319	\$6,703	Reading

Source: Common Core of Data, National Center for Education Statistics, retrieved October 5, 2005, from <http://nces.ed.gov/globallocator>. Per pupil expenditure are for 2002–2003 school year; all other indices are from 2003–2004 school year.

School Information

<i>General School Information</i>					<i>Student Demographics</i>					
School	District	Grade Range	Teacher FTEs	Students	Free/Reduced Lunch	Hispanic	African American	Asian American	White	American/Alaskan Indian
Middle School A	Athens	7–8	42.8	517	78.1%	0%	66.9%	4%	34.4%	0%
Middle School B	Bayou City	6–8	53.8	800	66.5%	1.1%	30.8%	0%	64.5%	3.6%
High School A	Bayou City	8–12	95.4	1,615	50.5%	0%	36%	0%	61.8%	1%
Elementary A	Highway Junction	PK–5	31.9	516	91%	2%	73%	0%	25.2%	0%
Elementary B	Highway Junction	PK–5	47	727	96.7%	0%	100%	0%	0%	0%

Source: Common Core of Data, National Center for Education Statistics, retrieved October 5, 2005, from <http://nces.ed.gov/globallocator>.

New Mexico Development Sites

District Information

District	Type Location	Schools	Teacher FTEs	Students	Expenditure per Pupil	Focus of SEDL's Work
Desert Hills	Urban Fringe of Large City	11	252.1	3,377	\$8,574	Reading
Farmville	Rural Inside MSA	6	100.9	1,514	\$7,406	Reading

Source: Common Core of Data, National Center for Education Statistics, retrieved October 5, 2005, from <http://nces.ed.gov/globallocator>. Per pupil expenditure are for 2002–2003 school year; all other indices are from 2003–2004 school year.

School Information

<i>General Information</i>					<i>Student Demographics</i>					
School	District	Grade Range	Teacher FTEs	Students	Free/Reduced Lunch	Hispanic	African American	Asian American	White	American/Alaskan Indian
Elementary A	Desert Hills	PK–5	13.9	149	76.5%	35.8%	0%	0%	4%	60.4%
Middle School A	Desert Hills	6–8	44.9	592	98%	60.8%	.5%	0%	11.1%	27.5%
Elementary B	Farmville	KG–5	12.5	194	94.8%	92.8%	0%	0%	7.2%	0%
Elementary C	Farmville	3–5	4	237	89%	84.8%	0%	0%	15.2%	0%
Elementary D	Farmville	PK–2	14.5	254	98.8%	87.8%	0%	0%	12.2%	0%
Middle School C	Farmville	6–8	22.8	397	87.9%	90.9%	0%	0%	9%	0%
High School B	Farmville	9–12	33.3	398	79.3%	88.2%	0%	0%	11.8%	0%

Source: Common Core of Data, National Center for Education Statistics, retrieved October 5, 2005, from <http://nces.ed.gov/globallocator>.

Oklahoma Development Sites

District Information

District	Type Location	Schools	Teachers	Students	Expenditure per Pupil	Focus of SEDL's Work
Wrightsville	Small town	8	267.2	4,279	\$5,835	Reading
Bricktown	Large Central City	27	1,189	19,356	\$5,493	Reading

Source: Common Core of Data, National Center for Education Statistics, retrieved October 5, 2005, from <http://nces.ed.gov/globallocator>. Per pupil expenditure are for 2002–2003 school year; all other indices are from 2003–2004 school year.

School Information

<i>General School Information</i>					<i>Student Demographics</i>					
School	District	Grade Range	Teachers	Students	Free/Reduced Lunch	Hispanic	African American	Asian American	White	American/Alaskan Indian
Elementary A	Wrightsville	PK–5	24	311	90.4%	43.1%	16.4%	1.6%	36.3%	2.6%
Middle School A	Bricktown	6–8	43	631	62.1%	9.8%	17.4%	2.5%	63%	7.1%
High School A	Bricktown	9–12	88.3	1,789	43.5%	12%	12%	3.7%	59.1%	4.9%

Source: Common Core of Data, National Center for Education Statistics, retrieved October 5, 2005, from <http://nces.ed.gov/globallocator>.

Test Sites*

District Information

District	Type Location	Schools	Teacher FTEs	Students	Expenditure per Pupil	Focus of SEDL's Work
Grisham, AR	Small Town	7	229	3,290	\$6,430	Math
Roydale, AR	Small Town	6	130	1,675	\$6,271	Reading
Brookesville, OK	Large Central City	4	78.3	1,192	\$5,602	Reading
Pineland, TX	Small Town	7	253.2	3,316	\$6, 948	Reading

Source: Common Core of Data, National Center for Education Statistics, retrieved October 5, 2005, from <http://nces.ed.gov/globallocator>. Per pupil expenditure are for 2002–2003 school year; all other indices are from 2003–2004 school year.

* *Note.* Since only one site analyzed in this report was located in Texas, demographic information for that site is included only in this table.

School Information

<i>General School Information</i>					<i>Student Demographics</i>					
School	District	Grade Range	Teacher FTEs	Students	Free/Reduced Lunch	Hispanic	African American	Asian American	White	American/Alaskan Indian
Elementary C	Grisham, AR	3–4	30	531	99.4%	2%	76.6%	0%	20.9%	0%
Intermediate A	Grisham, AR	5–6	33	490	99.4%	1%	80%	0%	20.8%	0%
Elementary D	Roydale, AR	1–3	29	408	99.3%	1%	77.7%	0%	21.3%	0%
High School C	Roydale, AR	10–12	24	339	99.1%	0%	73.2%	0%	26.5%	0%
Elementary B	Brookesville, OK	PK–K	19	168	98.2%	38%	26.2%	2.3%	20.8%	12.5%
Elementary C	Brookesville, OK	1–5	19	457	84.5%	30.6%	31.1%	0%	21.9%	16.2%
Middle School B	Brookesville, OK	6–8	21.5	279	85.3%	31.5%	33.3%	0%	23.3%	11.5%
High School B	Brookesville, OK	9–12	17	288	70.1%	29.5%	31.6%	0%	25.7%	12.9%
Elementary A	Pineland, TX	1–2	31.9	364	67.9%	29.1%	29.4%	0%	40.4%	0%
Elementary B	Pineland, TX	3–5	54	738	67.8%	31%	30.5%	1.5%	36.7%	0%
Elementary C	Pineland, TX	PK–1	30.6	432	77.3%	35.2%	33.3%	0%	31.5%	0%

Source: Common Core of Data, National Center for Education Statistics, retrieved October 5, 2005, from <http://nces.ed.gov/globallocator>.

**APPENDIX B:
DATA COLLECTION INSTRUMENTS**

Site Contact Records

The following pages contain screen prints from the relational database used for entry, storage, and use of site contract records.

Data Management System - Site Contact Record A

 New SCR
 Delete SCR
 Print SCR
 to LOG
 User ewaters

A CONTACT INFO

Log ID: _____
 Contact Date: _____
 Site: _____
 Contact Type: _____
 Contact Location: _____

Created on 8/30/2005 Last Modified 8/30/2005

A Recorder _____ **Title** _____

Source of Info. _____ **Time Spent** _____ cumulative hrs. over _____ day(s)

Resources Used During the Contact

B

Itemized Resources Used (any resource specified above must be entered here) New

Category	Resource Title	Location	Description (optional)	Rsrc. Disseminated?	Quantity
				<input type="checkbox"/>	
				<input type="checkbox"/>	
				<input type="checkbox"/>	

C

Stage of Work ?

SITE ENTRY

DATA SCAN

SYSTEM EXP.

PLANNING ACTION

TAKING ACTION/MON. RESULTS

RECYCLING FOR CONT. IMPROVEMENT

OTHER _____ ?

Framework Areas Addressed ?

<p>Levels</p> <p><input type="checkbox"/> Classroom</p> <p><input type="checkbox"/> School</p> <p><input type="checkbox"/> District</p> <p><input type="checkbox"/> State</p> <p><input type="checkbox"/> N/A</p>	<p>Competencies</p> <p><input type="checkbox"/> Collecting Interpreting and Using Data</p> <p><input type="checkbox"/> Creating Coherence</p> <p><input type="checkbox"/> Forging Alliances</p> <p><input type="checkbox"/> Building Capacity</p> <p><input type="checkbox"/> Promoting Innovation</p> <p><input type="checkbox"/> N/A</p>	<p>Components</p> <p><input type="checkbox"/> Standards</p> <p><input type="checkbox"/> Curriculum and Instruction</p> <p><input type="checkbox"/> Assessment</p> <p><input type="checkbox"/> Policy and Governance</p> <p><input type="checkbox"/> Professional Staff</p> <p><input type="checkbox"/> Resources</p> <p><input type="checkbox"/> Family and Community</p> <p><input type="checkbox"/> N/A</p>	<p>Principles</p> <p><input type="checkbox"/> Shared Vision</p> <p><input type="checkbox"/> Supportive Organizational Structure</p> <p><input type="checkbox"/> Challenging Curriculum and Engaged Student Learning</p> <p><input type="checkbox"/> Culture of Continuous Inquiry and Improvement</p> <p><input type="checkbox"/> Facilitative Leadership</p> <p><input type="checkbox"/> Supportive Rel. Btwn. System & Surroundings</p> <p><input type="checkbox"/> N/A</p>
--	---	--	--

PTLC Phases Addressed

<p>Cycle#</p> <p>_____</p>	<p>Phases Implemented</p> <p><input type="checkbox"/> Study</p> <p><input type="checkbox"/> Select</p> <p><input type="checkbox"/> Plan</p> <p><input type="checkbox"/> Implement</p> <p><input type="checkbox"/> Analyze</p> <p><input type="checkbox"/> Intervene</p>	<p>Other Cycle#</p> <p>_____</p>	<p>Other Phases Implemented</p> <p><input type="checkbox"/> Study</p> <p><input type="checkbox"/> Select</p> <p><input type="checkbox"/> Plan</p> <p><input type="checkbox"/> Implement</p> <p><input type="checkbox"/> Analyze</p> <p><input type="checkbox"/> Intervene</p>
-----------------------------------	--	---	--

Outcomes (in terms of Intended Purpose)	?
	A≡
	{a}
Outcomes (other than Intended Purpose; Unanticipated Outcomes, Positive or Negative)	A≡
	{a}
Reflections/Insights on Event in Relation to Intended Purposes (including Resources Used)	?
	A≡
	{a}
Reflections/Insights (in the larger, overall context of the work, SEDL Framework, Other Sites, etc.)	A≡
	{a}

To keep in mind for future contacts (anticipations, impressions, things to watch out for)		?
		A {a}
Participant Next Steps		A {a}
SEDL Next Steps/Additional Issues		A {a}
Data/Materials Collected (leave blank if "none")		?
Date	Name/Type of materials collected	New

Data/Materials Collected (leave blank if "none")	
Date	Name/Type of materials collected

Use of Protocols

- Which protocol did you use?
- Did you follow the prescribed instructions when you used it?
 Yes No (If yes, please skip to Item 6.)
- What changes did you make either to the protocol itself or how it was used?
- Why did you make those changes?
- Did those changes address whatever concern(s) you have?
- How well did the protocol help structure the activities you expected to accomplish during this recent contact?
- What, if any, modifications would you recommend based on your use of the protocol during this recent contact with the site? Why?

Data Management System - Site Contact Record B

New SCR Delete SCR Print SCR to LOG User kherbert

A **CONTACT INFO**
Log ID : 2841
Contact Date :
Site : TX.Nova
Contact Type :
Contact Location :
Created on 1/31/2006 Last Modified 1/31/2006

B Recorder kherbert Title
Text description of intended purpose ?
Public Private

B **Plans & Anticipations** Reset Text Formatting

C **Setting, Climate, Group Affect, & Context**

Actual Event (Major Activities w/sufficient detail, Purposes, Resources, Groupings, Roles, & Content Focus)

Fall 2003–2004 Interview Questions

SEDL Interviewer:	_____	Position Title:	_____
District Name:	_____	District- or School-Level Employee:	_____
School Name:	_____	Grades Taught:	_____
Educator's Name:	_____	Subject Taught:	_____

1. What are the priorities in your school?
2. What are the priorities in your district?
 - If not addressed, ask about priorities regarding reading/math.
3. What does the school expect of teachers?
 - How have expectations of teachers been communicated?
 - By whom?
4. What does the school expect of students?
 - How have expectations of students been communicated?
 - By whom?
5. How much coordination or alignment do you see between what is taught in your school(s) and the state standards?
 - How much coordination or alignment do you see in your grade level across schools?
 - How much coordination or alignment do you see between adjacent grades?
 - How could alignment be improved in your school/district?
6. Does the district supply the necessary instructional resources to teach the required curriculum?
 - If not, what were you/they unable to teach?
7. Do teachers in your school or district collaborate?
8. Do professional development topics focus on the goal of improving student learning?
 - If no, what is the focus?
 - Are teachers actively engaged during professional development sessions?
 - How does follow-up normally occur?
9. Is there anything I didn't ask that you think I should have asked or should know about?

Spring 2005 Interview Questions

SEDL Interviewer:	_____	Position Title:	_____
District Name:	_____	District- or School-Level Employee:	_____
School Name:	_____	Grades Taught:	_____
Educator's Name:	_____	Subject Taught:	_____
SEDL Teams:	_____	Date:	_____

1. Tell me about when your school (teacher/principal) / district (district administrators) first started working with SEDL. (If can't answer, have them refer to "you").
 - How has the work with SEDL changed over time?
 - Have these experiences made you think differently about your role as a teacher/principal/superintendent etc.? If so, how?

2. In your opinion, what did your school (teacher/principal)/district (district administrators) hope to gain from the work with SEDL? And did it? If not, why not?
 - What did you hope to gain from this work? And did you? If not, why not?

3. Now we're going to talk about the work SEDL's been doing with the PTLC, such as studying the standards, modeling and planning instruction, looking at student work, and so forth.
 - What has your school/district been doing to implement the PTLC?
 - How could the implementation of the PTLC be improved?
 - What are the two biggest benefits of using the PTLC?

4. Since your school/district started working with SEDL, what changes have you seen, if any, in the way teachers and administrators work together?

5. If SEDL were to do this project again, what should be done differently?
 - What work will continue, what won't continue? Why?

6. Is there anything else about the work with SEDL that we haven't talked about that you want to add?

***Working Systemically* Survey**

General Instructions:

Please read each statement carefully and indicate the extent to which you believe it is true. Using a pencil or a black or blue pen, please score each item. Please be sure to fill in the bubble completely with a dark mark. It is important that you correctly bubble in your responses so they are accurately scanned. Below are examples of correct and incorrect markings:

Correct			Incorrect		
					

Instructions for Page 2:

On this page it is important that you respond to all of the items. Please be sure to mark all roles (positions) and certifications that you have. This will enable us to examine the responses by group and will help us develop a more complete understanding of the systemic environment in your district.

Instructions for Pages 3–5:

When responding to the items on pages 2–4 please be sure to only mark only one response per item; items with more than one response marked will be treated as a skipped item. Please score each item on a scale of 1 to 6, with 1 indicating “Never True” and 6 indicating “Always True.”

Please keep in mind that there are no right or wrong responses for the statements. We are seeking **your** perception of how well these items describe your school/district. Please do not dwell on the items; your first instinct is probably your most accurate reflection.

Southwest Educational Development Laboratory

School Name _____

Background Information

All responses are confidential. Please do not put your name on this survey.

Respondent information

Gender: M F



[Site Name]

What position do you currently have in this school or district? (Please check all of the following that apply.)

- A Teacher
 - Elementary school
 - Middle School / Junior High
 - Senior High School
- B Teacher's Aide or Educational Aide
- C School Volunteer
- D School Administrator
- E District Administrator
- F District Support staff
- G Member of district leadership team
- H Member of campus leadership team
- I Parent of child enrolled in this school district
- J Community Representative

How many years have you worked in education?

Number of Years

<1 1-3 4-6 7-10 >10

- | | | | | | |
|--|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| Number of years in current position in this school or district | <input type="radio"/> A | <input type="radio"/> B | <input type="radio"/> C | <input type="radio"/> D | <input type="radio"/> E |
| Number of years worked all positions in this school | <input type="radio"/> A | <input type="radio"/> B | <input type="radio"/> C | <input type="radio"/> D | <input type="radio"/> E |
| Number of years worked all positions in this district | <input type="radio"/> A | <input type="radio"/> B | <input type="radio"/> C | <input type="radio"/> D | <input type="radio"/> E |
| Total number of years teaching/administration experience | <input type="radio"/> A | <input type="radio"/> B | <input type="radio"/> C | <input type="radio"/> D | <input type="radio"/> E |

What is the highest degree you have earned?

- A Less than high school
- B High school diploma or GED
- C Associate's degree
- D Bachelor's degree
- E Master's degree
- F Doctorate

What certification(s) do you have?

- A Elementary
- B Secondary
 - A Math
 - B Science
 - C Reading
 - D Social Studies
 - E English
 - F Other Specify: _____

Never True	Rarely True	Sometimes True	Usually True	Often True	Always True
1	2	3	4	5	6

Working Systemically Questionnaire

Please indicate the extent to which you agree or disagree with the following statements as they apply to your school/district.		1	2	3	4	5	6
1	The staff work collaboratively to maintain the alignment of instruction across grades.	①	②	③	④	⑤	⑥
2	Supplies essential for instruction are available in sufficient quantity.	①	②	③	④	⑤	⑥
3	Our school helps parents learn how to support their child's development of reading and mathematical skills.	①	②	③	④	⑤	⑥
4	Specific intervention plans are in place for students who do not meet expected proficiency levels.	①	②	③	④	⑤	⑥
5	Each student is expected to achieve high standards.	①	②	③	④	⑤	⑥
6	District and school leaders seek input from teachers to determine professional development plans.	①	②	③	④	⑤	⑥
7	Research-based instructional strategies are the focus of our professional development.	①	②	③	④	⑤	⑥
8	Our school creates an environment that encourages all students to learn.	①	②	③	④	⑤	⑥
9	Space exists in our school building for staff and others to work together.	①	②	③	④	⑤	⑥
10	Student assessment data are used to adjust school improvement plans.	①	②	③	④	⑤	⑥
11	The individuals responsible for implementing each part of the school plan are held accountable for seeing that those plans and objectives are carried out.	①	②	③	④	⑤	⑥
12	Both administrators and teachers receive help to interpret student assessment data.	①	②	③	④	⑤	⑥
13	Administrators ensure that teachers have necessary resources to deliver high quality instruction.	①	②	③	④	⑤	⑥
14	What teachers are doing in the classroom matches administrators' expectations for instruction.	①	②	③	④	⑤	⑥
15	District and school leaders create the time necessary for staff to use data to solve problems and plan collaboratively.	①	②	③	④	⑤	⑥
16	All children receive a high quality education in this school regardless of which teacher they have.	①	②	③	④	⑤	⑥
17	The purpose of professional development programs is to improve student learning.	①	②	③	④	⑤	⑥

Never True	Rarely True	Sometimes True	Usually True	Often True	Always True
1	2	3	4	5	6

Working Systemically Questionnaire

Please indicate the extent to which you agree or disagree with the following statements as they apply to your school/district.		1	2	3	4	5	6
18	There is a high level of trust among stakeholders (administrators, teachers, and parents) in our school that supports working together.	①	②	③	④	⑤	⑥
19	Our school provides each student what s/he needs to succeed.	①	②	③	④	⑤	⑥
20	Student assessments are used at all grade levels to improve instruction.	①	②	③	④	⑤	⑥
21	Administrators and teachers work together to develop a shared vision of how the school should best meet the needs of all students.	①	②	③	④	⑤	⑥
22	Connections are clear and consistent between what we teach and what we assess.	①	②	③	④	⑤	⑥
23	Our school provides sufficient time for professional development.	①	②	③	④	⑤	⑥
24	Our community engages in discussions about education programs in the district.	①	②	③	④	⑤	⑥
25	All of the instructional resources in this school are aligned with the academic goals of the district.	①	②	③	④	⑤	⑥
26	All school personnel have the same access to resources to support student learning.	①	②	③	④	⑤	⑥
27	Data are used to make decisions in our district and schools.	①	②	③	④	⑤	⑥
28	Teachers have the support they need to use student assessment data to plan instruction.	①	②	③	④	⑤	⑥
29	The school budget directly supports the school's priorities.	①	②	③	④	⑤	⑥
30	Administrators hold teachers accountable for student learning.	①	②	③	④	⑤	⑥
31	Administrators know what is to be taught and learned in each grade and subject area.	①	②	③	④	⑤	⑥
32	Our school's vision of high quality instruction is evident in our classrooms every day.	①	②	③	④	⑤	⑥
33	The principal visits each classroom in our school to monitor instruction.	①	②	③	④	⑤	⑥
34	Teachers have a voice in how resources are allocated.	①	②	③	④	⑤	⑥
35	Teachers and parents work together to improve student learning.	①	②	③	④	⑤	⑥

Never True	Rarely True	Sometimes True	Usually True	Often True	Always True
1	2	3	4	5	6

***Working Systemically* Questionnaire**

Please indicate the extent to which you agree or disagree with the following statements as they apply to your school/district.		1	2	3	4	5	6
36	Our school has the necessary resources to support research-based practices that promote student learning.	①	②	③	④	⑤	⑥
37	Teachers at each school meet regularly in grade-level or content-area meetings to discuss instructional issues.	①	②	③	④	⑤	⑥
38	Teachers know exactly what is to be taught and learned for each grade and subject area they teach.	①	②	③	④	⑤	⑥
39	Administrators expect teachers to participate in professional development programs to improve student learning.	①	②	③	④	⑤	⑥
40	Our school's success depends on students from each racial/ethnic groups performing well.	①	②	③	④	⑤	⑥
41	Our school routinely disaggregates student data to assess how well we're doing.	①	②	③	④	⑤	⑥
42	Administrators' actions demonstrate their commitment to improving student learning.	①	②	③	④	⑤	⑥
43	When students are promoted from one grade to the next, they have the prerequisite skills necessary to succeed at the next grade.	①	②	③	④	⑤	⑥
44	Our school's curriculum is closely aligned with the state standards.	①	②	③	④	⑤	⑥
45	The roles and responsibilities for achieving the school's goals and objectives are clearly documented in school plans.	①	②	③	④	⑤	⑥
46	Resources are reallocated based on reviews of student needs.	①	②	③	④	⑤	⑥
47	Our professional development programs match our school's priorities.	①	②	③	④	⑤	⑥
48	Our school's plan to improve student learning is aligned with the district's plan.	①	②	③	④	⑤	⑥
49	Professional development programs improve overall school performance.	①	②	③	④	⑤	⑥
50	The community and school share a common vision for their children.	①	②	③	④	⑤	⑥

**APPENDIX C:
DESCRIPTIONS OF STATE ASSESSMENT SYSTEMS**

Arkansas

The Arkansas Benchmark Tests are criterion-referenced exams developed around the Arkansas Curriculum Frameworks by teacher committees, the Arkansas Department of Education, and a test contractor. There are three tests administered to students at grades 4, 6, and 8 assessing literacy and mathematics achievement.

The following table provides the schools working with SEDL on the school improvement model in Arkansas and the tested grade level(s) at each school.

SEDL Schools and Tested Grade Levels: Arkansas

District	School	4 th	6 th	8 th
Forked River	Elementary School B (R)	X	X	
Grisham	Elementary School C (M)	X		
Grisham	Middle School A (M)		X	
Forked River	High School B (R)			X

Note. (M) = Math; (R) = Reading

The Arkansas Department of Education has defined four levels of student achievement for the benchmark exams (Arkansas Department of Education, 2004). These achievement levels are described below.

- **Advanced:** Advanced students demonstrate superior performance well beyond proficient grade-level performance. They can apply established reading, writing, and mathematics skills to solve complex problems and complete demanding tasks on their own. They can make insightful connections between abstract and concrete ideas and provide well-supported explanations and arguments.
- **Proficient:** Proficient students demonstrate solid academic performance for the grade tested and are well-prepared for the next level of schooling. They can use established reading, writing, and mathematics skills and knowledge to solve problems and complete tasks on their own. Students can tie ideas together and explain the ways their ideas are connected.
- **Basic:** Basic students show substantial skills in reading, writing, and mathematics; however, they only partially demonstrate the abilities to apply these skills.
- **Below Basic:** Below basic students fail to show sufficient mastering of skills in reading, writing, and mathematics to attain the basic level.

By 2004–2005, students who do not meet proficient or advanced status will be identified for remediation according to state guidelines (Arkansas Department of Education, 2005). Therefore, analysis of achievement levels focus on this minimum standard as the basis for comparisons.

Louisiana

The LEAP 21 and GEE 21 benchmark exams are criterion-referenced tests developed by the state to assess how well students have mastered the state content standards. The LEAP 21 tests are administered at grades 4 and 8 in mathematics, English language arts, science, and social studies. The Graduate Exit Examination is administered in grades 10 (English/language arts and math) and 11 (science and social studies). Since SEDL's work concentrates on reading and math development, student achievement data have been collected solely for the math and English language arts exams.

The table below highlights the schools working with SEDL on the school improvement model in Louisiana and tested grade levels at each school in math and English/language arts.

SEDL Schools and Tested Grade Levels: Louisiana

District	School	4 th	8 th	10 th
Highway Junction	Elementary School A (R)	X		
Highway Junction	Elementary School B (R)	X		
Athens	Middle School A (M)		X	
Bayou City	Middle School B (M)		X	
Bayou City	High School A (M)			X

Note. (M) = Math; (R) = Reading

The Louisiana Department of Education has identified five levels of student achievement for the LEAP 21 and GEE 21 benchmark exams (Louisiana Department of Education, 2004).

These achievement levels are described below.

- **Advanced:** A student at this level has demonstrated superior performance beyond the proficient level of mastery.
- **Mastery:** A student at this level has demonstrated competency over challenging subject matter and is well-prepared for the next level schooling.
- **Basic:** A student at this level has demonstrated only the fundamental knowledge and skills needed for the next level of schooling.
- **Approaching Basic:** A student at this level has only partially demonstrated the fundamental knowledge and skills needed for the next level of schooling.

- **Unsatisfactory:** A student at this level has not demonstrated the fundamental knowledge and skills needed for the next level of schooling.

For this analysis, results have been organized by percentage of students at or above the “Basic” level for fourth graders and percentage of students at or above the “Approaching Basic” level for eighth graders and tenth graders. The reason for this difference is that current state policy dictates that fourth-grade students must achieve “Basic” level status on at least one of the exams (math or English/language arts) in order to qualify for promotion in Louisiana schools. In eighth grade, students must attain “Approaching Basic” status on both exams in order to be eligible for promotion. Tenth-grade students must attain “Approaching Basic” status to pass a particular test. Graphs depict the percentage of students at or above “Basic” level for fourth graders and the percentage of students at or above “Approaching Basic” level for eighth and tenth graders as “% Meeting Standard.”

New Mexico

The New Mexico Statewide Articulated Assessment Program (NMSAAP) reflects a collection of criterion- and norm-referenced exams used by the state to measure student achievement. Standards-based assessments have only been recently adopted. However, since these are aligned with the state standards, the analysis of New Mexico achievement focuses on grade levels addressed by these exams, namely the New Mexico Standards-based Assessment (NMSA).

The New Mexico Standards-based Assessment is a set of criterion-referenced exams that reflect New Mexico benchmarks and standards. Through 2003–2004, these exams are administered at fourth and eighth grade in language arts and mathematics (New Mexico Department of Education, 2004). However, in 2004–2005, the number of tested grades increased, creating a discontinuity between reported achievement between the 2003–2004 and 2004–2005 school years. The basic issue is that New Mexico calculates Annual Yearly Progress (AYP) at the school level. However, since the testing structure changed between years, proficiency levels reported in 2003–2004 are not directly comparable to proficiency reports in 2004–2005 since the new calculations may include more tested students. For example, a middle school with grades 6, 7, and 8 would only have reported for proficiency in 2003–2004. The same school however would report proficiency levels in 2004–2005 that reflect the aggregate performance of all students in grades 6 through 8. For that reason, this section only reports on data that are directly comparable (i.e., 2002–2003 and 2003–2004).

The focus of the analyses is on standards-based assessment data because this parallels SEDL's work to improve alignment between content, curriculum and instruction, and assessment. Since the focus of SEDL's work is on reading content, data are presented for the standards-based assessment in reading.

The following table provides the schools working with SEDL on the school improvement model in New Mexico and the tested grade level(s) at each school on the standards-based assessment.

SEDL Schools and Tested Grade Levels: New Mexico

District	School	4th	8th
Desert Hills	Elementary School A (R)	X	
Farmville	Elementary School B (R)	X	
Farmville	Elementary School C (R)	X	
Desert Hills	Middle School A (R)		X
Farmville	Middle School C (R)		X

Note. (M) = Math; (R) = Reading

The New Mexico Public Department of Education has identified four performance-level categories for the standards-based assessment: advanced, proficient, nearing proficiency, and beginning step. However, in published reports, the department of education presents results in terms of students who achieve proficient status or better and those who do not.

Oklahoma

The Oklahoma Core Curriculum Tests (OCCT) are criterion-referenced exams linked to the Priority Academic Student Skills (PASS) or state curriculum. Currently, multiple choice tests are administered at grades 3, 4, 5, 7, and 8 in mathematics, reading, science, social studies, geography, U.S. history, and Constitution and government. It should be noted that not all subjects are tested at each grade, and the math tests in third and fourth grade are only field tests.

The focus of this section is on test data reflecting fifth- and eighth-grade reading achievement.³⁴ The table below describes the schools working with SEDL on the school improvement model in Oklahoma and the tested grade level(s) at each school for which we have data.

SEDL Schools and Tested Grade Levels: Oklahoma

District	School	5 th	8 th
Wrightsville	Elementary School A (R)	X	
Brookesville	Elementary School D (R)	X	
Bricktown	Middle School A (R)		X
Brookesville	Middle School B (R)		X

Note. (M) = Math; (R) = Reading

The Oklahoma Department of Education has defined four levels of student achievement for the benchmark exams for each subject and at each grade level (Oklahoma Department of Education, 2004). These performance descriptions are described below.

³⁴ The use of the literacy exam reflects the fact that SEDL's focus among these schools is solely on reading. Therefore, an examination of math results would be superfluous because SEDL has not been involved in any math activities.

Oklahoma Fifth-Grade Reading Performance Descriptors

- **Advanced:** Students consistently demonstrate a thorough understanding of the knowledge and skills expected of all students at this grade level. These skills are broadly demonstrated in reading processes, response to text, and acquisition of information through research. In addition to demonstrating a broad and in-depth understanding and application of all skills at the satisfactory performance level, students scoring at the advanced level typically use a wide range of strategies to interpret and evaluate text; regularly demonstrate a thorough and comprehensive understanding of literary forms; and consistently apply many different strategies for assessing, organizing, analyzing, synthesizing, and paraphrasing information.
- **Satisfactory:** Students demonstrate a general understanding of the reading knowledge and skills expected of all students at this grade level. Students scoring at the satisfactory level typically read and comprehend grade-level reading material using the following skills:
 - a. identify new words and use word meaning resources;
 - b. identify major elements of story structure;
 - c. recognize and interpret relationships in narrative and expository text;
 - d. determine central purpose, theme, or key concept/main idea and important details;
 - e. make inferences and draw conclusions/generalizations;
 - f. interpret figurative language and characteristics of poetry;
 - g. identify and analyze characteristics of a variety of genres;
 - h. distinguish among fact, supported inferences, and opinion in expository text;
 - i. determine author's purpose and point of view;
 - j. identify similarities and differences between reading selections; and
 - k. demonstrate use of functional print, informational resources, charts, and diagrams.
- **Limited Knowledge:** Students demonstrate a partial understanding of the reading knowledge and skills expected of all students at this grade level. Students scoring at the limited knowledge level are inconsistent in demonstrating satisfactory level competencies and typically demonstrate reading skills within more explicit and concrete contexts.
- **Unsatisfactory:** Students do not demonstrate at least a limited knowledge level of the skills expected of all students at this grade level. Students scoring at the unsatisfactory level should be given comprehensive reading instruction.

Oklahoma Eighth-Grade Reading Performance Descriptors

- **Advanced:** Students consistently demonstrate a thorough understanding of the knowledge and skills expected of all students at this grade level. These skills are broadly demonstrated in reading processes, response to text, and acquisition of information through research. In addition to demonstrating a broad and in-depth understanding and application of all skills at the satisfactory performance level, students scoring at the advanced level typically use a wide range of strategies to interpret and evaluate text; regularly demonstrate a thorough and comprehensive understanding of literary forms; and consistently apply many different strategies for assessing, organizing, analyzing, synthesizing, and paraphrasing information.
- **Satisfactory:** Students demonstrate a general understanding of the reading knowledge and skills expected of all students at this grade level. Students scoring at the satisfactory level typically read and comprehend grade-level reading material using the following skills:
 - a. determine literal and non-literal word meanings using a variety of strategies;
 - b. analyze informational text, poetry, short stories, novels, and dramas;
 - c. determine main idea and themes (stated or implied) and recognize relevance of details;
 - d. interpret figurative language and elements of poetry;
 - e. infer, predict, and generalize ideas;
 - f. judge author's purpose/point of view, accuracy of text, and fact/opinion; and
 - g. use appropriate strategies to organize and summarize information.
- **Limited Knowledge:** Students demonstrate a partial understanding of the reading knowledge and skills expected of all students at this grade level. Students scoring at the limited knowledge level are inconsistent in demonstrating the satisfactory level competencies and typically demonstrate reading skills within more explicit and concrete contexts.
- **Unsatisfactory:** Students do not demonstrate at least a limited knowledge level of the skills expected of all students at this grade level. Students scoring at the unsatisfactory level should be given comprehensive reading instruction.

For the analyses in this report, results have been organized by proportion of students meeting proficiency requirements for the state, currently defined as achieving at the “Satisfactory” level or above.

Texas

The Texas Assessment of Knowledge and Skills (TAKS) is actually an extension of the Texas Assessment of Academic Skills (TAAS) system immediately preceding it. Like the previous state testing systems discussed in this section, the TAKS is a criterion-referenced exam aligned with the Texas Essential Knowledge and Skills (TEKS) state curriculum.

In Spring 2003, Texas administered the TAKS for the first time in grades 3, 4, and 5. In addition to reading and math tests, fourth graders received a writing test. In fifth grade, a new science test was also administered. Due to the focus of the *Working Systemically* model on reading, the analysis in this section focuses on reading achievement for third-, fourth-, and fifth-graders at Elementary School A in the Pineland District.

The Texas State Board of Education has defined the performance standards for TAKS tests in terms of “Met Standard” (Texas Education Agency, 2004). The analysis examines the proportion of students meeting the standards in 2002–2003, 2003–2004, and 2004–2005 in the SEDL school as well as the matched-composite school.

**APPENDIX D:
INTERVIEW RATING RUBRICS**

Generic Rubric Template

Generic: A function related to working systemically is fulfilled in a site.				
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Respondent reports implementation failure or barriers to implementation. No structures or processes are reported to fulfill the function.	Respondent reports that structures (e.g., curriculum, meetings, or planned activities) or processes (e.g., communication) only partially fulfill the function. Educators are making minimal efforts to support the function.	Respondent reports that specific structures (e.g., curriculum, meetings, or planned activities) or processes (e.g., communication) to fulfill the function are in place.	Respondent reports regular implementation of the structures (e.g., curriculum, meetings, or planned activities) or processes (e.g., communication) to fulfill the function in some parts of the site.	Respondent reports that structures are fully implemented/in place throughout the site to support the model.

Expectations and Priorities Rubric

1. Expectations of teachers and school/district priorities are both focused on improving student learning in the SEDL focus subject area.				
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Respondent does not report any evidence of explicit priorities around student learning in the school or district.	Respondent reports that improving student learning is an explicit priority. Expectations of teachers are not explicitly focused on student learning priorities.	Respondent reports that stated expectations of teachers and school/district priorities are both focused on specific student learning goals.	Respondent reports that educators in some parts of the site are expected to engage in activities or processes that are consistent with school/district priorities (e.g., reviewing student data and making adjustments in instruction).	Respondent reports that expectations of teachers and school priorities are aligned to support student learning across all levels (e.g., levels of administration, principals, and teachers) of the site. Behaviors, activities, and policies of the site are linked by this common goal.

Alignment Rubric

2. There is alignment between curriculum and the state standards in the SEDL focus subject area.				
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Respondent reports that there is no written curriculum aligned to state standards.	Respondent reports development of a written curriculum aligned to standards. The site is reported to have a partially aligned curriculum in the SEDL focus subject area.	Respondent reports that there is a written curriculum that is aligned to state standards in the SEDL focus area.	Respondent reports that an aligned written curriculum is used vertically and/or horizontally within some parts of the site (e.g., within some grade levels or schools but not others).	Respondent reports consistent use of an aligned written curriculum vertically and horizontally across SEDL schools in the site.

Collaboration Rubric

3. Educators collaborate to improve student learning in the SEDL focus subject area outside of professional development sessions.				
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Respondent does not report that educators collaborate, and/or collaboration is couched negatively.	Respondent reports infrequent collaboration (e.g., hallway meetings, lunchtime). No report of a clear focus on student learning. Educators are reported to be making efforts to collaborate.	Respondent reports that formal collaborative meetings focused on improving student learning in the SEDL focus area have taken place (e.g., grade-level meetings, department meetings, campus leadership teams).	Respondent reports that educators regularly schedule formal meetings in some parts of the site (e.g., some grade levels or schools but not others) in which educators are focused on improving student learning (e.g., planning instruction, developing curriculum).	Respondent reports that both teachers and school/district administrators are engaged in ongoing collaboration throughout the site (e.g., cross-school collaborative meetings; administrator meetings across schools; superintendent attends cross-school meetings and/or supports formal collaborative processes).

Professional Development Rubric

4. Educators use school- or district-sponsored professional development to improve student learning in the SEDL focus subject area.				
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Respondent does not report that professional development promotes skills related to student learning.	Respondent reports that some professional development activities support skills related to student learning. Professional development does not appear to be intentionally focused on student learning goals or driven by student learning needs in the site.	Respondent reports that there is a common, intentional professional development focus on building skills related to student learning needs and objectives.	Respondent reports that educators regularly schedule professional development sessions to support skills related to student learning needs in some parts of the site (e.g., some grade levels, subjects, or groups of teachers). Follow-up is reported to occur but is inconsistent or irregular.	Respondent reports that all professional development is driven by the site's student learning needs and goals. Follow-up/monitoring is reported to occur consistently throughout the site.

**APPENDIX E:
COMPLETE DATA SOURCES**

Site Contact Records (SCRs)

Overview of Site Contact Record Entries

Site	SCRs Begin	SCRs End	Total Entries
Farmville, NM	4/18/01	4/27/05	148
Piedmont, AR	4/16/01	1/10/03	179
Southwest City, TX	5/9/01	6/1/03	53
Desert Hills, NM	5/15/01	5/4/05	179
Cottonwood, NM	5/16/01	5/13/03	108
Mesa, NM	6/24/01	10/26/04	97
Highway Junction, LA	8/8/01	4/20/05	163
Deforest, TX	8/8/01	10/8/02	48
Delta Village, AR	8/9/01	8/13/04	173
Wrightsville, OK	8/10/01	5/26/05	88
Tornillo, TX	8/21/01	6/5/02	23
Oak Hill, AR	8/27/01	9/26/02	40
Bricktown, OK	9/18/01	5/12/05	146
Washington City, TX	9/28/01	10/13/03	69
High Meadows, OK	10/11/01	5/24/04	84
River City, LA	11/26/01	4/15/04	65
Forked River, AR	12/5/01	5/16/05	166
Athens, LA	3/20/02	5/10/05	77
Bayou City, LA	8/27/02	5/4/05	74
Brookesville, OK	1/13/03	5/11/05	128
Pineland, TX	1/30/03	5/13/05	41
Grisham, AR	3/4/03	1/24/05	43
Roydale, AR	3/31/03	5/4/05	110
Total			2,302

Interviews

Educators Interviewed: Fall

	Respondents	<i>Fall 2002</i>		<i>Fall 2003</i>		<i>Fall 2004</i>	
		<i>N</i>	<i>Pct</i>	<i>N</i>	<i>Pct</i>	<i>N</i>	<i>Pct</i>
<i>District</i>	Superintendents and Assistant Superintendents	10	11.1%	20	16.7%	13	11.7%
	Curriculum Specialists	10	11.1%	9	7.5%	12	10.8%
	Other	7	7.8%	2	1.7%	6	5.4%
	<i>Total District-Level Interviews</i>	27	30.0%	31	25.9%	31	27.9%
<i>School</i>	Principals and Assistant Principals	20	22.2%	32	26.7%	29	26.1%
	Teachers	39	43.3%	55	45.8%	42	37.8%
	Other	4	4.4%	2	1.7%	10	9.0%
	<i>Total School-Level Interviews</i>	63	69.9%	89	74.2%	81	72.9%
	Total Interviews	90	100%	120	100%	111	100%
	Total Sites	12		17		15	

Note. Percentages are rounded to whole numbers.

Educators Interviewed: Spring

	Respondents	<i>Spring 2003</i>		<i>Spring 2004</i>		<i>Spring 2005</i>	
		<i>N</i>	<i>Pct</i>	<i>N</i>	<i>Pct</i>	<i>N</i>	<i>Pct</i>
<i>District</i>	Superintendents and Assistant Superintendents	10	15.9%	15	13.6%	14	15.2%
	Curriculum Specialists	6	9.5%	9	8.2%	8	8.7%
	Other	8	12.7%	10	9.1%	4	4.3%
	<i>Total District-Level Interviews</i>	24	38.1%	34	30.9%	26	28.3%
<i>School</i>	Principals and Assistant Principals	20	31.7%	26	23.6%	21	22.8%
	Teachers	17	27.0%	47	42.7%	36	39.1%
	Other	2	3.2%	3	2.7%	9	9.8%
	<i>Total School-Level Interviews</i>	39	61.9%	76	69.1%	66	71.7%
	Total Interviews	63	100%	110	100%	92	100%
	Total Sites	15		15		12	

Note. Percentages are rounded to whole numbers.

Working Systemically Survey

Working Systemically Survey Respondents

	Respondents	2002		2003		2004		2005	
		N	Pct	N	Pct	N	Pct	N	Pct
<i>District</i>	Administrators	3	0.3%	27	3.6%	19	2.5%	23	3.0%
	Other	0	0.0%	14	1.9%	11	1.4%	7	0.9%
	<i>Total District-Level Surveys</i>	3	0.3%	41	5.5%	30	3.9%	30	3.9%
<i>School</i>	Administrators	34	3.9%	30	4.0%	42	5.4%	45	5.9%
	Teachers	682	77.8%	633	84.9%	621	80.4%	480	63.5%
	Other	84	9.6%	18	2.4%	49	6.3%	29	3.8%
	<i>Total School-Level Surveys</i>	800	91.2%	681	91.3%	712	92.2%	554	73.3%
	Unknown	74	8.4%	24	3.2%	30	3.9%	172	22.8%
	Total Respondents	877	100%	746	100%	772	100%	756	100%
	Total Sites	17	100%	12	100%	12	100%	12	100%

Note. Six sites provide survey responses across all four spring administrations.

Student Achievement Data

The most consistent achievement data across sites starts in 2002–2003. Prior to this period, only three states have relevant data. Additionally, not all sites started working with SEDL at the same time; therefore, baseline years are variable even within states. Starting analyses with 2003 gives us a relatively uniform dataset through 2004 and 2005 (when available).

Student Achievement Data by State

State	2001	2002	2003	2004	2005
Arkansas	X	X	X	X	
Louisiana	X	X	X	X	X
New Mexico			X	X	
Oklahoma		X	X	X	
Texas			X	X	X

**APPENDIX F:
CHI-SQUARE ANALYSES**

The following sections present contingency tables and chi-square (χ^2) statistics for grade-level analyses presented in Section V. Statistics are presented after each table. For 2 x 2 contingency tables, chi-square statistics indicate whether two groups differ. In tables with more than four cells (e.g., 2 x 3), chi-square investigates whether there are significant differences between groups in the table. For example, in a 2 x 3 table there are three possible comparisons: group 1 vs. group 2, group 1 vs. group 3, and group 2 vs. group 3. A significant result suggests a difference may exist between one or more pairs of groups. In these cases, multiple comparisons are conducted among proportions through pairwise contrasts where χ^2_{ψ} is the chi-square value for each contrast. Results for these comparisons are presented when appropriate. In some instances, no schools or only one school matched the SEDL school on demographic variables and initial achievement. When no schools matched, researchers used the schools remaining after demographic selection criteria were applied. Comparison groups are labeled as No Matched Comparison (NMC) or Single School where appropriate.

Arkansas

Grisham

Grade 4 Math, AR, Grisham–Elementary School C

SEDL

	2003	2004	Row Totals
Proficient	72	95	167
Basic and Below	140	162	302
Column Totals	212	257	469

$\chi^2 = .46 < 3.84 = .95 \chi_1^2, p > .05$

Comparison Group (No Matched Comparison Group on Initial Achievement)

	2003	2004	Row Totals
Proficient	30	40	70
Basic and Below	37	40	77
Column Totals	67	80	147

$\chi^2 = .40 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Proficient	72	30	95	40	237
Basic and Below	140	37	162	40	379
Column Totals	212	67	257	80	616

$\chi^2 = 7.68 < 7.82 = .95 \chi_3^2, p > .05$

Grade 6 Math, AR, Grisham–Middle School A

SEDL

	2003	2004	Row Totals
Proficient	10	23	33
Basic and Below	198	206	404
Column Totals	208	229	437

$\chi^2 = 4.28 > 3.84 = .95 \chi_1^2, p < .05$

Comparison Group

	2003	2004	Row Totals
Proficient	10	14	24
Basic and Below	174	114	288
Column Totals	184	128	312

$\chi^2 = 3.22 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Proficient	10	10	23	14	57
Basic and Below	198	174	206	114	692
Column Totals	208	184	229	128	749

$\chi^2 = 7.51 < 7.82 = .95 \chi_3^2, p > .05$

Forked River

Grade 4 Literacy, AR, Forked River-Elementary School B

SEDL

	2003	2004	Row Totals
Proficient	19	16	35
Basic and Below	25	22	47
Column Totals	44	38	82

$\chi^2 = .01 < 3.84 = .95 \chi_1^2, p > .05$

Comparison Group

	2003	2004	Row Totals
Proficient	26	30	56
Basic and Below	29	31	60
Column Totals	55	61	116

$\chi^2 = .04 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Proficient	19	26	16	30	91
Basic and Below	25	29	22	31	107
Column Totals	44	55	38	61	198

$\chi^2 = .66 < 7.82 = .95 \chi_3^2, p > .05$

Grade 6 Literacy, AR, Forked River–Elementary School B

SEDL

	2003	2004	Row Totals
Proficient	14	20	34
Basic and Below	27	31	58
Column Totals	41	51	92

$\chi^2 = .25 < 3.84 = .95 \chi_1^2, p > .05$

Comparison (Single School)

	2003	2004	Row Totals
Proficient	10	16	26
Basic and Below	20	21	41
Column Totals	30	37	67

$\chi^2 = .69 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Proficient	14	10	20	16	60
Basic and Below	27	20	31	21	99
Column Totals	41	30	51	37	159

$\chi^2 = 1.00 < 7.82 = .95 \chi_3^2, p > .05$

Grade 8 Literacy, AR, Forked River–High School B

SEDL

	2003	2004	Row Totals
Proficient	10	22	32
Basic and Below	31	35	66
Column Totals	41	57	98

$\chi^2 = 2.19 < 3.84 = .95 \chi_1^2, p > .05$

Comparison Group (No Matched Comparison Found on Initial Achievement)

	2003	2004	Row Totals
Proficient	37	43	80
Basic and Below	25	37	62
Column Totals	62	80	142

$\chi^2 = .50 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Proficient	10	37	22	43	112
Basic and Below	31	25	35	37	128
Column Totals	41	62	57	80	240

$\chi^2 = 15.50 > 7.82 = .95 \chi^2_{3}, p < .05$

Pairwise Contrasts for SEDL vs. Comparison Group

	χ^2_{ψ}	$.95 \chi^2_{3}$	Sig.
SEDL 2003 vs. Comparison Group 2003	14.86	7.82	$p < .05$
SEDL 2004 vs. Comparison Group 2004	3.16	7.82	$p > .05$

Louisiana

Athens

Grade 8 Math, LA, Athens–Middle School A

SEDL

	2003	2004	2005	Row Totals
Approaching Basic and Above	145	153	141	439
Unsatisfactory	57	44	72	173
Column Totals	202	197	213	612

$\chi^2 = 6.64 > 5.99 = .95 \chi^2, p < .05$

Pairwise Contrasts for Multiple Years of Testing in Athens–Middle School A

	χ^2_{ψ}	$.95 \chi^2$	Sig.
SEDL 2003 vs. SEDL 2004	1.84	5.99	$p > .05$
SEDL 2004 vs. SEDL 2005	6.81	5.99	$p < .05$
SEDL 2003 vs. SEDL 2005	1.52	5.99	$p > .05$

Comparison Group

	2003	2004	2005	Row Totals
Approaching Basic and Above	79	99	93	271
Unsatisfactory	36	21	31	88
Column Totals	115	120	124	359

$\chi^2 = 6.07 > 5.99 = .95 \chi^2, p < .05$

Pairwise Contrasts for Multiple Years of Testing in Comparison Group

	χ^2_{ψ}	$.95 \chi^2$	Sig.
Comparison 2003 vs. Comparison 2004	6.20	5.99	$p < .05$
Comparison 2004 vs. Comparison 2005	2.07	5.99	$p > .05$
Comparison 2003 vs. Comparison 2005	1.18	5.99	$p > .05$

SEDL vs. Comparison Group within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	SEDL 2005	Comp. 2005	Row Totals
Approaching Basic and Above	145	79	153	99	141	93	710
Unsatisfactory	57	36	44	21	72	31	261
Column Totals	202	115	197	120	213	124	971

$\chi^2 = 14.19 > 11.07 = .95 \chi^2_{5^2}, p < .05$

Pairwise Contrasts for SEDL vs. Comparison Group

	χ^2_{ψ}	$.95 \chi^2_{2^2}$	Sig.
Comparison 2003 vs. Comparison 2004	.332	11.07	$p > .05$
Comparison 2004 vs. Comparison 2005	1.12	11.07	$p > .05$
Comparison 2003 vs. Comparison 2005	3.02	11.07	$p > .05$

Bayou City

Grade 8 Math, LA, Bayou City–Middle School B

SEDL

	2003	2004	2005	Row Totals
Approaching Basic and Above	180	165	149	494
Unsatisfactory	73	66	105	244
Column Totals	253	231	254	738

$\chi^2 = 11.99 > 5.99 = .95 \chi^2_{2^2}, p < .05$

Pairwise Contrasts for Multiple Years of Testing in Athens–Middle School A

	χ^2_{ψ}	$.95 \chi^2_{2^2}$	Sig.
SEDL 2003 vs. SEDL 2004	.00	5.99	$p > .05$
SEDL 2004 vs. SEDL 2005	8.87	5.99	$p < .05$
SEDL 2003 vs. SEDL 2005	8.83	5.99	$p < .05$

Comparison Group (No Matched Comparison Found on Initial Achievement)

	2003	2004	2005	Row Totals
Approaching Basic and Above	155	149	161	465
Unsatisfactory	27	21	30	78
Column Totals	182	170	191	543

$\chi^2 = .87 < 5.99 = .95 \chi^2_{2^2}, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	SEDL 2005	Comp. 2005	Row Totals
Approaching Basic and Above	180	155	165	149	149	161	959
Unsatisfactory	73	27	66	21	105	30	322
Column Totals	253	182	231	170	254	191	1,281

$\chi^2 = 72.79 > 11.07 = .95 \chi^2_5, p < .05$

Pairwise Contrasts for SEDL vs. Comparison Group

	χ^2_ψ	$.95 \chi^2_6$	Sig.
SEDL 2003 vs. Comparison 2003	13.05	11.07	$p < .05$
SEDL 2004 vs. Comparison 2004	17.30	11.07	$p < .05$
SEDL 2005 vs. Comparison 2005	39.87	11.07	$p < .05$

Grade 10 Math, LA, Bayou City–High School A

SEDL

	2003	2004	2005	Row Totals
Approaching Basic and Above	234	249	222	705
Unsatisfactory	91	85	89	265
Column Totals	325	334	311	970

$\chi^2 = .93 < 5.99 = .95 \chi^2_2, p > .05$

Comparison Group

	2003	2004	2005	Row Totals
Approaching Basic and Above	273	236	241	750
Unsatisfactory	124	74	83	281
Column Totals	397	310	324	1,031

$\chi^2 = 5.40 < 5.99 = .95 \chi^2_2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	SEDL 2005	Comp. 2005	Row Totals
Approaching Basic and Above	234	273	249	236	222	241	1,455
Unsatisfactory	91	124	85	74	89	83	546
Column Totals	325	397	334	310	311	324	2,001

$\chi^2 = 6.33 < 11.07 = .95 \chi^2_5, p > .05$

Highway Junction

Grade 4 English/Language Arts, LA, Highway Junction–Elementary School A

SEDL

	2003	2004	2005	Row Totals
Basic and Above	41	40	41	122
Approaching Basic and Below	30	36	34	100
Column Totals	71	76	75	222

$\chi^2 = .39 < 5.99 = .95 \chi^2, p < .05$

Comparison Group

	2003	2004	2005	Row Totals
Basic and Above	42	35	39	116
Approaching Basic and Below	33	31	22	86
Column Totals	75	66	61	202

$\chi^2 = 1.64 < 5.99 = .95 \chi^2, p < .05$

SEDL vs. Comparison Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	SEDL 2005	Comp. 2005	Row Totals
Basic and Above	41	42	40	35	41	39	238
Approaching Basic	30	33	36	31	34	22	186
Column Totals	71	75	76	66	75	61	424

$\chi^2 = 2.29 < 11.07 = .95 \chi^2, p > .05$

Grade 4 English/Language Arts, Highway Junction–Elementary School B

SEDL

	2003	2004	2005	Row Totals
Basic and Above	79	89	59	227
Approaching Basic and Below	32	37	34	103
Column Totals	111	126	93	330

$\chi^2 = 1.73 < 5.99 = .95 \chi^2, p < .05$

Comparison (Single School)

	2003	2004	2005	Row Totals
Basic and Above	76	73	65	214
Approaching Basic and Below	34	26	19	79
Column Totals	110	99	84	293

$\chi^2 = 1.70 < 5.99 = .95 \chi_2^2, p > .05$

SEDL vs. Comparison School Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	SEDL 2005	Comp. 2005	Row Totals
Basic and Above	79	76	89	73	59	65	441
Approaching Basic	32	34	37	26	34	19	182
Column Totals	111	110	126	99	93	84	623

$\chi^2 = 4.77 < 11.07 = .95 \chi_5^2, p > .05$

New Mexico

Desert Hills

Grade 4 Reading, NM, Desert Hills–Elementary School A

SEDL

	2003	2004	Row Totals
Proficient or Better	6	10	16
Nearing Proficiency or Below	11	15	26
Column Totals	17	25	42

$\chi^2 = .10 < 3.84 = .95 \chi_1^2, p > .05$

Comparison Group

	2003	2004	Row Totals
Composite			
Proficient or Better	11	12	23
Nearing Proficiency or Below	22	17	39
Column Totals	33	29	62

$\chi^2 = .43 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Proficient or Better	6	11	10	12	39
Nearing Proficiency or Below	11	22	15	17	65
Column Totals	17	33	25	29	104

$\chi^2 = .53 < 7.82 = .95 \chi_3^2, p > .05$

Grade 8 Reading, NM, Desert Hills–Middle School A

SEDL

	2003	2004	Row Totals
Proficient or Better	77	93	170
Nearing Proficiency or Below	101	100	201
Column Totals	178	193	371

$\chi^2 = .91 < 3.84 = .95 \chi_1^2, p > .05$

Comparison Group

	2003	2004	Row Totals
Proficient or Better	99	110	209
Nearing Proficiency or Below	123	100	223
Column Totals	222	210	432

$\chi^2 = 2.62 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Proficient or Better	77	99	93	110	379
Nearing Proficiency or Below	101	123	100	100	424
Column Totals	178	222	193	210	803

$\chi^2 = 4.05 < 7.82 = .95 \chi_3^2, p > .05$

Farmville

Grade 4 Reading, NM, Farmville–Elementary School B

SEDL

	2003	2004	Row Totals
Proficient or Better	8	11	19
Nearing Proficiency or Below	27	28	55
Column Totals	35	39	74

$\chi^2 = .28 < 3.84 = .95 \chi_1^2, p > .05$

Comparison Group

	2003	2004	Row Totals
Proficient or Better	10	9	19
Nearing Proficiency or Below	31	27	58
Column Totals	41	36	77

$\chi^2 = .00 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Proficient or Better	8	10	11	9	38
Nearing Proficiency or Below	27	31	28	27	113
Column Totals	35	41	39	36	151

$\chi^2 = .30 < 7.82 = .95 \chi_3^2, p > .05$

Grade 4 Reading, NM, Farmville–Elementary School C

SEDL

	2003	2004	Row Totals
Proficient or Better	22	26	48
Nearing Proficiency or Below	51	40	91
Column Totals	73	66	139

$\chi^2 = 1.31 < 3.84 = .95 \chi_1^2, p > .05$

Comparison Group

	2003	2004	Row Totals
Proficient or Better	11	12	23
Nearing Proficiency or Below	26	23	49
Column Totals	37	35	72

$\chi^2 = .17 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Proficient or Better	22	11	26	12	71
Nearing Proficiency or Below	51	26	40	23	140
Column Totals	73	37	66	35	211

$\chi^2 = 1.64 < 7.82 = .95 \chi_3^2, p > .05$

Grade 8 Reading, NM, Farmville–Middle School C

SEDL

	2003	2004	Row Totals
Proficient or Better	37	50	87
Nearing Proficiency or Below	78	80	158
Column Totals	115	130	245

$\chi^2 = 1.05 < 3.84 = .95 \chi_1^2, p > .05$

Comparison (Single School)

	2003	2004	Row Totals
Proficient or Better	41	67	108
Nearing Proficiency or Below	70	58	128
Column Totals	111	125	236

$\chi^2 = 6.58 > 3.84 = .95 \chi_1^2, p < .05$

SEDL vs. Comparison School Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Proficient or Better	37	41	50	67	195
Nearing Proficiency or Below	78	70	80	58	286
Column Totals	115	111	130	125	481

$\chi^2 = 13.01 > 7.82 = .95 \chi_3^2, p < .05$

Pairwise Contrasts for SEDL vs. Comparison School

	χ^2_{ψ}	$.95 \chi_3^2$	Sig.
SEDL 2003 vs. Comparison Group 2003	.57	7.82	$p > .05$
SEDL 2004 vs. Comparison Group 2004	6.01	7.82	$p > .05$

Oklahoma

Bricktown

Grade 8 Reading, OK, Bricktown–Middle School A

SEDL

	2003	2004	Row Totals
Satisfactory or Above	130	128	258
Limited Knowledge or Below	51	38	89
Column Totals	181	166	347

$\chi^2 = 1.27 < 3.84 = .95 \chi_1^2, p > .05$

Comparison Group

	2003	2004	Row Totals
Satisfactory or Above	134	150	284
Limited Knowledge or Below	51	38	89
Column Totals	185	188	373

$\chi^2 = 2.78 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Satisfactory or Above	130	134	128	150	542
Limited Knowledge or Below	51	51	38	38	178
Column Totals	181	185	166	188	720

$\chi^2 = 4.32 < 7.82 = .95 \chi_3^2, p > .05$

Brookesville

Grade 5 Reading, OK, Brookesville–Elementary School D

SEDL

	2003	2004	Row Totals
Satisfactory or Above	25	27	52
Limited Knowledge or Below	39	42	81
Column Totals	64	69	133

$\chi^2 = .00 < 3.84 = .95 \chi_1^2, p > .05$

Comparison Group

	2003	2004	Row Totals
Satisfactory or Above	21	21	42
Limited Knowledge or Below	30	16	46
Column Totals	51	37	88

$\chi^2 = 2.09 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Satisfactory or Above	25	21	27	21	94
Limited Knowledge or Below	39	30	42	16	127
Column Totals	64	51	69	37	221

$\chi^2 = 3.74 < 7.82 = .95 \chi_3^2, p > .05$

Grade 8 Reading, OK, Brooksville–Middle School B

SEDL

	2003	2004	Row Totals
Satisfactory or Above	39	30	69
Limited Knowledge or Below	24	23	47
Column Totals	63	53	116

$\chi^2 = .34 < 3.84 = .95 \chi_1^2, p > .05$

Comparison Group

	2003	2004	Row Totals
Satisfactory or Above	19	17	36
Limited Knowledge or Below	11	8	19
Column Totals	30	25	55

$\chi^2 = .13 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Satisfactory or Above	39	25	30	18	112
Limited Knowledge or Below	24	13	23	10	70
Column Totals	63	38	53	28	182

$\chi^2 = .93 < 7.82 = .95 \chi_3^2, p > .05$

Wrightsville

Grade 5 Reading, OK, Wrightsville–Elementary School B

SEDL

	2003	2004	Row Totals
Satisfactory or Above	19	29	48
Limited Knowledge or Below	17	4	21
Column Totals	36	33	69

$\chi^2 = 10.02 > 3.84 = .95 \chi_1^2, p < .05$

Comparison Group

	2003	2004	Row Totals
Satisfactory or Above	20	19	39
Limited Knowledge or Below	19	14	33
Column Totals	39	33	72

$\chi^2 = .29 < 3.84 = .95 \chi_1^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	Row Totals
Satisfactory or Above	19	20	29	19	87
Limited Knowledge or Below	17	19	4	14	54
Column Totals	36	39	33	33	141

$\chi^2 = 12.81 > 7.82 = .95 \chi_3^2, p < .05$

Pairwise Contrasts for SEDL vs. Comparison Group

	χ^2_{ψ}	$.95 \chi_3^2$	Sig.
SEDL 2003 vs. Comparison Group 2003	.02	7.82	$p > .05$
SEDL 2004 vs. Comparison Group 2004	8.64	7.82	$p < .05$

Texas

Pineland

Grade 3 Reading, TX, Pineland–Elementary School A

SEDL

	2003	2004	2005	Row Totals
Met Standard	192	206	193	591
Did Not Meet Standard	37	31	17	85
Column Totals	229	237	210	676

$\chi^2 = 6.56 > 5.99 = .95 \chi^2, p < .05$

Pairwise Contrasts for Multiple Years of Testing in Pineland–Elementary School A, Grade 3 Reading

	χ^2_{ψ}	$.95 \chi^2$	Sig.
SEDL 2003 vs. SEDL 2004	.88	5.99	$p > .05$
SEDL 2004 vs. SEDL 2005	2.98	5.99	$p > .05$
SEDL 2003 vs. SEDL 2005	6.87	5.99	$p < .05$

Comparison Group

	2003	2004	2005	Row Totals
Met Standard	85	86	97	268
Did Not Meet Standard	14	11	7	32
Column Totals	99	97	104	300

$\chi^2 = 2.99 < 5.99 = .95 \chi^2, p > .05$

SEDL vs. Comparison Group within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	SEDL 2005	Comp. 2005	Row Totals
Met Standard	192	85	206	86	193	97	859
Did Not Meet Standard	37	14	31	11	17	7	117
Column Totals	229	99	237	97	210	104	976

$\chi^2 = 10.25 < 11.07 = .95 \chi^2, p > .05$

Grade 4 Reading, TX, Pineland–Elementary School A

SEDL

	2003	2004	2005	Row Totals
Met Standard	166	150	157	473
Did Not Meet Standard	41	42	75	158
Column Totals	207	192	232	631

$\chi^2 = 10.61 > 5.99 = .95 \chi^2, p < .05$

Pairwise Contrasts for Multiple Years of Testing in Pineland–Elementary School A, Grade 4 Reading

	χ^2_{ψ}	$.95 \chi^2$	Sig.
SEDL 2003 vs. SEDL 2004	.26	5.99	$p > .05$
SEDL 2004 vs. SEDL 2005	5.96	5.99	$p > .05$
SEDL 2003 vs. SEDL 2005	9.17	5.99	$p < .05$

Comparison Group

	2003	2004	2005	Row Totals
Met Standard	84	91	75	250
Did Not Meet Standard	20	16	21	57
Column Totals	104	107	96	307

$\chi^2 = 1.65 < 5.99 = .95 \chi^2, p > .05$

SEDL vs. Comparison Group Within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	SEDL 2005	Comp. 2005	Row Totals
Met Standard	166	84	150	91	157	75	723
Did Not Meet Standard	41	20	42	16	75	21	215
Column Totals	207	104	192	107	232	96	938

$\chi^2 = 17.58 > 11.07 = .95 \chi^2, p < .05$

Pairwise Contrasts for SEDL vs. Comparison Group

	χ^2_{ψ}	$.95 \chi^2$	Sig.
SEDL 2003 vs. Comparison 2003	.01	11.07	$p > .05$
SEDL 2004 vs. Comparison 2004	2.30	11.07	$p > .05$
SEDL 2005 vs. Comparison 2005	4.01	11.07	$p > .05$

Grade 5 Reading, TX, Pineland–Elementary School A

SEDL

	2003	2004	2005	Row Totals
Met Standard	160	157	160	477
Did Not Meet Standard	42	31	31	104
Column Totals	202	188	191	581

$\chi^2 = 1.77 < 5.99 = .95 \chi^2_{2}, p > .05$

Comparison Group

	2003	2004	2005	Row Totals
Met Standard	88	80	86	254
Did Not Meet Standard	26	13	17	56
Column Totals	114	93	103	310

$\chi^2 = 2.95 < 5.99 = .95 \chi^2_{2}, p > .05$

SEDL vs. Comparison Group within Year

	SEDL 2003	Comp. 2003	SEDL 2004	Comp. 2004	SEDL 2005	Comp. 2005	Row Totals
Met Standard	160	88	157	80	160	86	731
Did Not Meet Standard	42	26	31	13	31	17	160
Column Totals	202	114	188	93	191	103	891

$\chi^2 = 4.73 < 11.07 = .95 \chi^2_{5}, p > .05$