

NATIONAL WRITING PROJECT

LOCAL SITE RESEARCH INITIATIVE REPORT

Cohort I (2003-2004)

FEBRUARY 2005

ACKNOWLEDGEMENTS

This summary report was written by Judy Buchanan, Elyse Eidman-Aadahl, Linda Friedrich, Paul LeMahieu, and Richard Sterling and edited by Roxanne Barber and Judith Bess, all of whom are National Writing Project staff. Jolene Galegher played the role of consulting research editor and, in that capacity, edited each research report and prepared the written summaries that form the basis for this report.

Members of four writing project sites along with their university colleagues designed and conducted the studies that provide the evidentiary and analytic basis for the first cohort of the Local Site Research Initiative. We acknowledge their hard work as well as the work of the many teachers at each site who assisted with scoring student work and conducting field work. The research teams are:

Mississippi Writing/Thinking Institute, Mississippi State University, Starkville, MS

Principal investigator: Sherry Seale Swain

Research team: David Morse, Richard L. Graves, Linda R. Irby, Robert Harris, Linda Allsup, Mandy Goldman, Emily Noble, Tanisha Rushing, Kim Patterson, Lois Connington, Lyneshia Griffin, and LaShanda Miller

National Writing Project at Kent State University, Kent, OH

Principal investigators: Nancy McCracken and David Bruce

Research team: Anthony Manna, Judith Sewell, and Deb DeBenedictis

Pennsylvania Writing and Literature Project, West Chester University, West Chester, PA

Principal investigators: Mary Buckelew and Andrea Fishman

Research team: Diane Dougherty, Krishna Kumar, Vicki Steinberg, Lynne Dorfman, Rose Capelli, and Dori Hoch

South Coast Writing Project, University of California, Santa Barbara, CA

Principal investigators: Sheridan D. Blau, Rosemary H. Cabe, Nada Rayyes, and Anne Whitney

Research team: Michelle Woodbridge, Tim Dewar, Chia-Chen Lai, Paul Rogers, and Sarah Boggs

Finally, special thanks are due to the many teachers and thousands of students who participated in these studies and gave time and effort to working with the local research teams.

National Writing Project
February 2005

National Writing Project • University of California • 2105 Bancroft Way #1042
Berkeley, CA 94720 • 510-642-0963 • www.writingproject.org

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	OVERVIEW OF NWP LOCAL SITE RESEARCH INITIATIVE (LSRI)	1
	A. Proposal Selection Process.....	1
	B. Summary of LSRI Key Learnings.....	2
III.	SUMMARY OF LSRI STUDIES AND RESULTS	3
	A. Program Emphasis	3
	B. Grade Levels Examined.....	3
	C. Research Approaches	3
	D. Table: Local Site Research Initiative Summary of Results by Site	4
IV.	SUMMARY ANALYSIS OF LSRI STUDIES	10
	A. Mississippi Writing/Thinking Institute, Mississippi State University <i>Analysis of the Effect of a Multi-year District Partnership: Fourth-Grade Level</i>	10
	B. Mississippi Writing/Thinking Institute, Mississippi State University <i>Analysis of the Effect of Multi-year School Partnerships: Seventh-Grade Level</i>	10
	C. Pennsylvania Writing and Literature Project, West Chester University <i>Analysis of the Effect of the First Year of a Professional Development Program</i>	12
	D. South Coast Writing Project, University of California at Santa Barbara <i>Analysis of the Effect of a Mature Professional Development Program on a Variety of Instructional and School Improvement Goals</i>	13
	E. National Writing Project at Kent State University <i>Analysis of the Effects on Teachers' Practices and Student Performance of Teachers' Participation in the Invitational Summer Institute</i>	14
V.	LESSONS LEARNED – ANALYSIS OF LIMITATIONS AND BENEFITS OF THE LSRI APPROACH	15
	A. Analysis of Methodological and Programmatic Limitations.....	15
	B. Analysis of Benefits of the LSRI Approach.....	18
VI.	FUTURE DIRECTIONS – DESIGNING A NATIONAL STUDY	20
VII.	REFERENCES	21

NATIONAL WRITING PROJECT

FEBRUARY 2005

LOCAL SITE RESEARCH INITIATIVE REPORT – Cohort I (2003-2004)

I. INTRODUCTION

This report describes the activities and presents the results of the first year of the Local Site Research Initiative (LSRI) of the National Writing Project (NWP). NWP structured the LSRI as a portfolio of locally based research efforts designed to study the effectiveness of the writing project model and its effects on students in a range of local contexts. During this first year, NWP also successfully designed and piloted a structure for the support of these local research efforts. Five local studies were conducted by four writing project sites in diverse settings: Mississippi Writing/Thinking Institute (Mississippi State University, MS), Pennsylvania Writing and Literature Project (West Chester University, PA), Northeast Ohio Writing Project at Kent State (Kent State University, OH), and South Coast Writing Project (University of California, Santa Barbara, CA). In this report, we share the results of these local studies and identify important learnings that inform the second cohort of LSRI studies currently being conducted.

II. OVERVIEW OF NWP LOCAL SITE RESEARCH INITIATIVE (LSRI)

The LSRI is intended to build a portfolio of rigorous, context-specific evaluation studies of local writing project work. A central feature of the NWP model is that the specific design of professional development programs is intended to vary according to local needs, reform priorities, and school conditions. Writing project sites emphasize common core principles of effective instruction, yet the design and delivery of services is negotiated with local education authorities. The goal of the LSRI is to develop an evaluation plan that can provide a group of studies that examine local practices based on these core principles and will illuminate teacher practices and student achievement in writing across a range of grade levels, schools, and local contexts.

A. Proposal Selection Process for LSRI Cohort I

Writing project sites were invited to participate through a request for proposal process and those selected were awarded \$20,000 and technical assistance from the national office. Twenty-three writing project sites submitted proposals for research to be conducted during the 2003-04 academic year. All proposals were reviewed both by NWP staff and by external researchers serving as peer reviewers. Reviewers evaluated the proposals for focus, content, and methodological rigor and appropriateness, as well as for feasibility and the capacity of the site to successfully execute its research plan.

All of the research studies were required to assess student writing performance using direct assessments and to incorporate some form of comparative reference to enable attribution of observed results. The approaches to research in each of the funded sites were designed to address the *Principles for Scientific Research in Education* developed by the National Research Council (National Research Council 2002).

Sites selected for initial funding were chosen both for the quality of their research plans and their ability to contribute to a collective portfolio that is complementary with respect to both the questions asked and the methods used to explore them. At the core, each site study was designed to investigate how some key component of their local program activities contributed to changes in teaching practices and to student learning and achievement in writing. The studies employed various kinds of data to examine student learning, including performance on state and local assessments, responses to writing prompts and/or revision tasks, and analyses of student portfolios.

Finally, NWP designed the LSRI technical assistance program to build the research capacity of local writing project sites. This capacity-building aspect of the LSRI is intended to contribute to the establishment of stronger research-based local programming in the districts and schools writing projects serve.

B. Summary of LSRI Key Learnings from the Pilot Year

- As expected, local program offerings required diverse research designs to capture appropriately program effects. Key reasons for this diversity were the need to adapt local programs to state standards, local reform initiatives, school or district interests, and state or district testing programs. Nonetheless, professional development inservice designs reflected a common focus on core practices in the teaching of writing across the contexts.
- With technical assistance and support, all four writing project sites were able to conduct studies of local programming that attended to the effect of teacher participation in professional development programs based on principles advocated by NWP on student performance through direct assessments of student writing, as well to the success of teacher implementation of core practices in the teaching of writing.
- While producing data that are moderately favorable to NWP local programming, methodological and administrative challenges led to unevenness in the research outcomes.
- An analysis of the experience of this first year pointed to clear methodological improvements needed, which have been incorporated into the second cohort of LSRI studies currently underway.

III. SUMMARY OF LSRI STUDIES AND RESULTS

Four of the five studies included in the first cohort analyzed the effect of writing project professional development inservice program offerings to schools and one study analyzed the effect of the writing project Invitational Summer Institute. The studies were conducted in diverse regions of the country and included teachers and students in rural, urban, and suburban areas. The studies also examined students with diverse economic, language, and racial and ethnic backgrounds. Taken together, these studies were intended to capture data related to teacher participation in writing project programs of varying intensity from participation in a first year of a professional development program to a retrospective study of the impact of a multi-year district partnership.

A. Program Emphasis

Four studies—the Mississippi Writing/Thinking Institute’s (MWTI) Fourth and Seventh Grade Studies, the Pennsylvania Writing and Literature Project’s (PAWLP) study, and the South Coast Writing Project’s (SCWriP) study—examined inservice work offered by writing project teacher-consultants to schools or districts in their local service areas. The MWTI studies examined the impact of ongoing, multi-year inservice programs developed and offered with the support of districts and schools respectively. The PAWLP study examined a one-year, districtwide professional development project that took place during the year of the study. The SCWriP study examined the impact of inservice programming on English Language Arts teachers who had participated for three years in a professional development program, but were no longer active participants. The Northeast Ohio Writing Project (NOWP) at Kent State study focused on the effect of the Invitational Summer Institute and participation in ongoing writing project professional development.

B. Grade Levels Examined

Two studies, the MWTI Fourth Grade Study and the PAWLP, focused on elementary school students. Two studies, the MWTI Seventh Grade and SCWriP studies, focused on middle school students. The NOWP study included students from grades 5–12.

C. Research Approaches

Central to each of the five studies was a commitment to understand what difference writing project professional development makes for teachers’ practice and, in turn, what difference teachers’ instructional practice makes for student learning. Each study, as required by the LSRI guidelines, employed direct assessments of student writing performance and each included some form of comparative reference.

Two studies, the MWTI Fourth and Seventh Grade Studies, compare multiple years of cross-sectional data drawn from the state-wide Mississippi Writing Assessment. Both studies compare the writing scores of students in districts or schools where MWTI has provided intensive inservice with demographically similar districts or schools as well as the state as a whole. Data in both studies are disaggregated by free/reduced lunch status, race/ethnicity, ELL, gender, and special education.

Each of the remaining three studies employs a quasi-experimental design and analyzes pre-/post-writing samples. The largest of these studies, the PAWLP, compares the growth of students in a rural district where it is providing intensive inservice programs with student growth in a demographically similar district. NOWP and SCWriP match individual teachers and classrooms. The PAWLP and NOWP at Kent State studies based their prompts, rubrics, and testing and scoring procedures on Pennsylvania and Ohio writing assessment systems, respectively. The SCWriP study employed the widely used Six Traits rubric and developed prompts similar to those used in several state writing assessments.

**D. Table: Local Site Research Initiative
Summary of Results by Site**

Study	Research Design			Results	Methodological and Program Implementation Challenges
	Outcome Studied	Measures	Samples		
<p>MS—Mississippi Writing/Thinking Institute (MWTI), Mississippi State University</p> <p><i>Analysis of Impact of a Multi-year District Partnership: Fourth Grade Level</i></p> <p><u>Program Emphasis</u> Three-year district – wide partnership at the elementary level</p> <p><u>Research Approach</u> Analysis of five years of successive panels of Mississippi Writing Assessment data</p>	<p><u>Teacher Outcomes</u> Degree of implementation of teaching practices modeled in MWTI professional development</p> <p><u>Student Outcomes</u> Achievement on a direct assessment of writing in comparison with students in demographically similar districts</p>	<p><u>Teacher Outcome Measures</u></p> <ul style="list-style-type: none"> • Interviews and classroom observations administered once annually to teachers in the program district • A four-point implementation continuum <p><u>Student Outcome Measures</u></p> <ul style="list-style-type: none"> • Mississippi Writing Assessment that uses a four-point holistic rubric to evaluate narrative or informative writing 	<p><u>Grade 4</u></p> <p><u>Program</u> Teachers = 44 Students = 1,138</p> <p><u>Comparison</u> Teachers = 40 Students = 1,070</p>	<p><u>Teacher Outcomes</u></p> <ul style="list-style-type: none"> • Teachers involved in the professional development increasingly employ the demonstrated teaching practices over time. • Teachers involved in the professional development increase the amount of time they spend on writing in their classrooms. <p><u>Student Outcomes</u></p> <ul style="list-style-type: none"> • Students of teachers involved in the professional development score significantly higher on the state writing assessment than do their counterparts in comparison districts. These findings are consistent over all three years of the program. • There is a positive correlation between the level at which teachers implement the writing project strategies and the writing assessment scores of their students. 	<ul style="list-style-type: none"> • This study relies on successive panels of student data. Therefore, writing assessments do not measure student growth.

Study	Research Design			Results	Methodological and Program Implementation Challenges
	Outcome Studied	Measures	Samples		
<p>MS—Mississippi Writing/Thinking Institute (MWTI), Mississippi State University</p> <p><i>Analysis of the Effect of a Multi-year School Partnership: Seventh Grade Level</i></p> <p><u>Program Emphasis</u> Multi-year middle school partnerships</p> <p><u>Research Approach</u> Analysis of two years of successive panels of seventh grade, Mississippi Writing Assessment data</p>	<p><u>Teacher Outcomes</u> Degree of implementation of teaching practices modeled in MWTI professional development</p> <p><u>Student Outcomes</u> Achievement on direct assessment of writing in comparison with students in demographically similar schools</p>	<p><u>Teacher Outcome Measures</u></p> <ul style="list-style-type: none"> Interviews and classroom observations administered once annually to teachers in the program schools A four-point implementation continuum <p><u>Student Outcome Measures</u></p> <ul style="list-style-type: none"> Mississippi Writing Assessment that uses a four-point holistic rubric to evaluate narrative or informative writing 	<p><u>Grade 7¹</u></p> <p><u>Program</u> Teachers = 10 Students = 510</p> <p><u>Comparison</u> Teachers = 8 Students = 776</p>	<p><u>Teacher Outcomes</u></p> <ul style="list-style-type: none"> Teachers involved in the professional development increasingly employ the demonstrated teaching practices over time. <p><u>Student Outcomes</u></p> <ul style="list-style-type: none"> The impact on seventh-grade student writing as measured by state writing assessments was uneven—in the first full year of the program, program students outscored their counterparts in comparison schools to a statistically significant degree. When there was less intensive professional development in the second year of the program, no statistically significant differences were seen between program and comparison students. 	<ul style="list-style-type: none"> This study, like the MWTI fourth-grade study, relies on successive panels of student data. Therefore, writing assessments do not measure student growth. In the second year of the program, 6 of the 10 participating teachers were new to the school, while only one teacher in the comparison sample turned over. Further, program teachers participated in fewer hours of MWTI professional development than they did in the preceding year. This reduced program intensity, particularly in interaction with the inclusion of teachers new to the programming, could account for the differing results from year to year.

¹ The numbers of teachers and students in program and comparison schools are based on 2001-02 data.

Study	Research Design			Results	Methodological and Program Implementation Challenges
	Outcome Studied	Measures	Samples		
<p>PA—Pennsylvania Writing and Literature Project (PAWLP), West Chester University</p> <p><i>Analysis of the Effect of the First Year of a Professional Development Program</i></p> <p><u>Program Emphasis</u> First year of an in-service partnership in a small rural district focused on elementary grades</p> <p><u>Research Approach</u> Quasi-experimental design, analyzing student growth on measures modeled after the Pennsylvania System of School Assessment (PSSA) Writing Test</p>	<p><u>Teacher Outcomes</u> Teachers’ implementation of writing instruction strategies modeled by PAWLP</p> <p><u>Student Outcomes</u> Growth on a direct assessment of student writing over the course of one year, comparing all students in a district where PAWLP provided inservice with students in a demographically similar district</p>	<p><u>Teacher Outcome Measures</u></p> <ul style="list-style-type: none"> • Observations • Semi-structured interviews <p><u>Student Outcome Measures</u></p> <ul style="list-style-type: none"> • Prompts and rubrics modeled after the PSSA writing assessment for grades 3–6 • PAWLP Domain Scoring Guide for K-2, developed and piloted by team of primary specialists and modeled after PSSA writing assessment 	<p><u>Grades K–5</u></p> <p><u>Program</u> Teachers = 39 Students = 837</p> <p><u>Comparison</u> Teachers = 35 Students = 794</p>	<p><u>Teacher Outcomes</u></p> <ul style="list-style-type: none"> • K–2 teachers exhibited high degree of implementation of recommended practices in their classrooms; 3–5 teachers made fewer changes because preparation for state reading and mathematics assessments reduced their ability to participate in professional development. <p><u>Student Outcomes</u></p> <ul style="list-style-type: none"> • Program students’ writing scores increased more than comparison group scores at all grade spans. These differences were statistically significant at K, 1, and 2, but not statistically significant in 3–5. 	<ul style="list-style-type: none"> • Sample attrition among the upper elementary grade program teachers

Study	Research Design			Results	Methodological and Program Implementation Challenges
	Outcome Studied	Measures	Samples		
<p>CA—South Coast Writing Project (SCWriP), University of California, Santa Barbara</p> <p><i>Analysis of the Effect of a Mature Professional Development Program on a Variety of Instructional and School Improvement Goals</i></p> <p><u>Program Emphasis</u> Three-year, middle-school partnership</p> <p><u>Research Approach</u> Quasi-experimental design, analyzing student growth on prompted timed writing tasks and Writing Apprehension Tests²</p>	<p><u>Teacher Outcomes</u> Program teachers' implementation of classroom practices modeled by SCWriP, in comparison with non-participating teachers</p> <p><u>Student Outcomes</u> Comparison of middle-school students of program teachers and students of matched comparison teachers from demographically similar schools:</p> <ul style="list-style-type: none"> • Growth on a prompted writing task administered at two points in time. • Changes in student attitudes toward writing 	<p><u>Teacher Outcomes</u></p> <ul style="list-style-type: none"> • Survey, interview protocol, and observation that examine responses to professional development and impact on classroom practice <p><u>Student Outcomes</u></p> <ul style="list-style-type: none"> • Writing prompts drawn from released items from state tests and scored using the Six-Trait scoring rubric³ • Writing Apprehension Test 	<p><u>Grades 6–8</u></p> <p><u>Program</u> Teachers = 4 Students = 170</p> <p><u>Comparison</u> Teachers = 4 Students = 176</p>	<p><u>Teacher Outcomes</u></p> <ul style="list-style-type: none"> • Teachers who had participated in the program were more likely than other teachers to encourage their students to think of writing as a collaborative process and to design exercises and assignments that called for various kinds of interactions among students and between students and the teacher. <p><u>Student Outcomes</u></p> <ul style="list-style-type: none"> • No statistically significant differences between students in the program and comparison groups on the pre-/post-writing sample. • Attitudes toward writing were similar across program and comparison groups. 	<ul style="list-style-type: none"> • Pretest was administered in December, three months after the beginning of the school year, and therefore may not provide a true baseline picture of students' writing achievement. • Lack of alignment between writing assessment and program goals • Attrition in the sample of both program and comparison teachers, and, therefore, in the student sample

² John A. Daly and Michael D. Miller, "The empirical development of an instrument to measure writing apprehension," *Research in the Teaching of English*, 9 (1975): 242-249.

³ Ruth Culham, *6+1 Traits of Writing* (New York: Scholastic Professional Books, 2003).

Study	Research Design			Results	Methodological and Program Implementation Challenges
	Outcome Studied	Measures	Samples		
<p>OH—National Writing Project at Kent State (NOWP), Kent State University</p> <p><i>Analysis of the Effects on Teachers’ Practices and Student Performance of Teachers’ Participation in the Invitational Summer Institute</i></p> <p><u>Program Emphasis</u> Invitational Summer Institute and ongoing writing project participation</p> <p><u>Research Approach</u> Quasi-experimental design, analyzing student growth on timed prompted writing assessments modeled after the Ohio Proficiency and Ohio Graduation writing tests</p>	<p><u>Teacher Outcomes</u> Comparing NOWP and non-NOWP teachers’ practices, with an emphasis on approach to teaching revision and the range of genres taught</p> <p><u>Student Outcomes</u> Comparing students of writing project teachers with students of teachers who have participated in other high quality literacy professional development:</p> <ul style="list-style-type: none"> • Growth on prompted writing tasks administered at two points in time • Use and knowledge of revision strategies 	<p><u>Teacher Outcomes</u></p> <ul style="list-style-type: none"> • Semi-structured interviews • Teacher recommendations to students on how to revise first draft writing • Content analysis of the types of writing included in student portfolios <p><u>Student Outcomes</u></p> <ul style="list-style-type: none"> • Prompts modeled after the Ohio Proficiency writing test and six-point holistic rubric drawn from the writing portion of the Ohio Graduation Test. • Proficiency sample essay revision task for students; scoring rubric based on the Ohio Graduation Test. 	<p><u>Grades 5–12</u></p> <p><u>Program</u> Teachers = 15 Students = 361</p> <p><u>Comparison</u> Teachers = 14 Students = 277</p>	<p><u>Teacher Outcomes</u></p> <ul style="list-style-type: none"> • Program teachers’ comments regarding revision to student writing focused more on issues such as audience than on mechanical issues than did non-NOWP teachers <p><u>Student Outcomes</u></p> <ul style="list-style-type: none"> • Students in program classrooms scored higher than students in comparison classrooms during the pre-test, the revision task, and the post-test writing. • In the semester-long high school classes, where pre-test was administered at the beginning of instruction, students in program classrooms improve their writing from pre-test to post-test more than students in comparison classrooms. 	<ul style="list-style-type: none"> • Pre-test was administered in December in grades 5–8, three months after the beginning of the school year, and therefore may not provide a true picture of students’ writing ability prior to instruction. Therefore, the observed differences in student performance could be due either to a selection artifact or to differential impact of instruction during the first three months of the year.

IV. SUMMARY ANALYSIS OF LSRI STUDIES

A. Mississippi Writing/Thinking Institute, Mississippi State University – *Analysis of the Effect of a Multi-year District Partnership: Fourth-Grade Level*

This study, by Mississippi Writing/Thinking Institute (MWTI), investigated impact in a district with a several year history of involvement with MWTI and modest teacher turnover. Taken together, this district provided a window into the potential impact of a mature professional development program.

This study focuses on fourth-grade students and reports results over a period of five years, including two years that are treated as baseline performance and three years in which post-intervention results are examined. Teacher practices and student performance were examined in the Program District and in two smaller, but demographically similar, Comparison Districts. Analyses of teacher practices rely, primarily, on interviews with the teachers and school and classroom observations by researchers; scores on state writing assessments were used as measures of student performance. It is important to note that, as this multi-year program focused on a single grade, the students in question differ across program years.

The results of this investigation indicate that teachers in the Program District employ MWTI-advocated practices to a greater extent as they gain experience with the program. The percentage of teachers in the Program District that named MWTI strategies as part of their teaching strategies increased from 55% in the first post-intervention year to 93% in the third post-intervention year. A similar pattern was observed in the percentage of teachers that reported that they used writing as a tool for learning in their classrooms and in the amount of instructional time teachers devoted to writing in their classrooms. Overall, teacher implementation and student results both favored the Program District. The scores on state writing assessments of students of teachers involved in MWTI professional development programs, relative to those of students in the Comparison Districts, were significantly better in all three post-intervention years. Further, the degree of implementation of MWTI-advocated practices was correlated with the degree of improvement in student performance; the greater the extent to which teachers implement the MWTI techniques, the more student performance improved.

In short, this study yielded two important results regarding the effects of participation in MWTI programs by fourth-grade teachers. First, the results indicate that exposure to MWTI professional development over time increases the likelihood that MWTI-practices will be implemented. Second, the greater the extent to which MWTI-advocated practices are implemented, the more positive the effect on student performance. In sum, MWTI programs—and likely those at other NWP sites—should be treated as long-term enterprises in which the likelihood and the magnitude of the desired “payoff” increases as teachers’ knowledge of and experience with the MWTI-advocated practices increases.

B. Mississippi Writing/Thinking Institute, Mississippi State University –
Analysis of the Effect of Multi-year School Partnerships: Seventh-Grade Level

This investigation focuses on seventh grade students and was conducted over a two-year period in two sets of schools, referred to here as Program Schools and Comparison Schools, in Mississippi. Although the duration and level of involvement of these schools in the activities of the Mississippi Writing/Thinking Institute (MWTI) varied, all schools had received professional development services for at least two years prior to the study.

The study was undertaken to determine whether the teachers in these schools apply the teaching strategies presented in the MWTI professional development programs and how their students perform on state writing assessments compared to students in demographically similar schools whose teachers have not participated in MWTI programs. During the two-year period of the study, as well as in preceding years during which some MWTI programs had been conducted, there were 10 Language Arts teachers in the set of Program Schools and another 8 teachers in the Comparison Schools. As this multi-year program focused on a single grade, the students in question differ across program years.

Analyses of teacher practices, intended to assess implementation of core practices advocated during professional development, involve only teachers in the Program Schools. They rely on interviews with the teachers, as well as a 30-minute observation of the teacher's classroom. All data regarding teachers' practices were collected during the second intervention year. Among the teachers whose practices were examined in this analysis, self-reported rates of implementation of MWTI-advocated practices were very high, with almost 100% of teachers reporting that they had implemented key practices, which included prewriting, editing, production of multiple drafts, peer response, teacher conferences, publishing, presentation of mini-lessons, and revision. Eight of the 10 teachers reported spending 50% or more of class time on writing or writing-related activities, and a majority of the teachers were able to articulate ways in which they use assessment results in the classroom, as well as ways they expand their expertise as writing teachers.

Analyses of student performance indicated that during the first intervention year, 2002–2003, the proportion of students who obtained a score of 2 or better on a 4-point scale was the same for the two sets of schools, but a significantly higher proportion of students (38%) in Program Schools obtained scores of 3 or better than in Comparison Schools (25%). Further, in the first intervention year, scores on the state writing assessment were significantly higher among students in the Program Schools than for students in the state as a whole. This difference, however, did not hold up in the second year of the intervention. In that year, the proportion of students who obtained scores of 3 or better in the Program Schools was 54%, whereas in the Comparison Schools, the proportion was 58%. In the second intervention year, scores of students in the Program Schools were statistically indistinguishable from those of students in the state as a whole.

The most important question to arise from this study is why the positive results of teachers' participation in MWTI programs observed in the first intervention year did not occur in the second year. As the study of student outcomes for fourth-graders whose teachers participated in MWTI professional development programs included in this report indicates, several years of professional development services may be required to reach high levels of implementation of the MWTI-advocated practices and resulting student performance. The authors report that in this study, teacher participation was less intensive during the second intervention year as schools juggled time allocations for professional development and teaching. There were fewer opportunities for teachers to participate in MWTI's professional development activities, the structure of those activities changed from a focus on workshops and classroom demonstrations to a focus on coaching and mentoring, and, in two of the Program Schools, professional development time focused on content-area teaching rather than on writing and language arts. Further, the rate of teacher turnover in the Program Schools was 60%, meaning that six of the 10 Language Arts teachers in these schools were new to the school during the second intervention year, and experienced only the less intensive professional development activities. In the Comparison Schools, turnover was only 10%. It is possible that these changes in professional development activities and in personnel account for the different patterns of performance observed in the two intervention years.

C. Pennsylvania Writing and Literature Project, West Chester University – *Analysis of the Effect of the First Year of a Professional Development Program*

In this investigation, the writing project took advantage of an invitation to present a program of professional development for teachers in several schools in a rural district in southeast Pennsylvania. The invitation was issued by district administrators, in anticipation of changes in the Pennsylvania state assessment of writing. The professional development program consisted of a complex mix of elements including full-day workshops, study group sessions, opportunities to observe expert teachers, and classroom coaching.

The study investigated the effect of the program on teacher practice and sought to document any resulting effect on student performance in the first year. Assessments included students' scores on standardized tests and observations of and interviews with teachers. The study focused on a program school and used a comparison design. With support of the district, a suitable comparison school was identified and agreed to participate in the study based on the expectation that it would be allowed access to the professional development in the following year.

For younger students and their teachers, the results on direct assessment of student writing were encouraging. In kindergarten, as well as grades 1 and 2, end-of-year scores for students whose teachers had participated in the professional development program were higher and had increased at a greater rate than those whose teachers had not participated. In addition, interviews and observations documented a high degree of implementation of recommended practices in program teachers' classrooms. In grades 3–5, the same pattern of stronger increase in performance for the NWP program participants

was observed. However, the differences were not statistically significant at the given sample sizes. In other words, at all grade levels, students whose teachers participated in the professional development program performed better than students whose teachers did not participate. The differences were statistically significant at grades K–2 but not at grades 3–5.

The authors reported a number of difficulties that affected the sample and analysis of grades 3–5, most notably the withdrawal from the professional development program of four teachers at the fifth grade. This reduction in the sample size, motivated by the need of these teachers to direct their efforts towards assessment in math and reading, make the results at the upper grades inconclusive.

**D. South Coast Writing Project, University of California, Santa Barbara –
*Analysis of the Effect of a Mature Professional Development Program on a Variety of Instructional and School Improvement Goals***

This study examined the effects of participation in IIMPaC, a professional development program based on NWP principles developed and offered by the South Coast Writing Project, based at the University of California, Santa Barbara. IIMPaC was designed to help middle school teachers work effectively with students in low-performing public schools in which a high proportion of students come from families in which English is not their home language. Intended outcomes of the program include adoption of teacher practices associated with high academic achievement in writing and the promotion of a positive professional development culture associated with ongoing school reform and improvement. This study was conducted within a comparison design emphasizing matched pairs of teachers working with similar students and in similar schools.

To assess teacher change, the investigators studied multiple aspects of teacher behavior using interviews, observations, and surveys; they also evaluated both student performance on timed writing tasks—one offered in December and the other offered in June—and asked students to report their attitudes toward writing on a well-known instrument called the Writing Apprehension Test (Daly and Miller 1975). The rubric used to assess student writing was drawn from the widely-used 6 Trait system (Culham 2003). The teacher participants were four middle school teachers who had participated in the IIMPaC program and four who had not; the student participants were members of classes taught by these teachers.

The results indicate that participation in IIMPaC had substantial effects on how teachers conceive of themselves as teachers and writers and on their teaching practice. Although all of the teachers had been exposed to some postgraduate instruction in the teaching of writing, those who had participated in IIMPaC were more likely to think of themselves as part of a collaborative professional community of teachers than were teachers who had not participated in this program. This difference—reflected in the kinds of interactions teachers have with each other and with their students—is consistent with NWP principles regarding writing as a communicative and intellectual process requiring research and revision. In addition to viewing themselves as members of a community of writers,

teachers who had participated in IIMPaC were more likely than other teachers to encourage their students to think of writing as a collaborative process and to design exercises and assignments that called for various kinds of interactions among students and between students and the teacher. Teachers also reported an increased sense of professionalism and collaboration.

An analysis of student performance, however, did not yield statistically significant differences between the program students and the comparison students. Further, the attitudes toward writing of the two groups of students were similar. Scores on the Writing Apprehension Test did not differ across groups.

The researchers surmise that the measure of performance used to assess student performance may have been insufficiently sensitive to the kinds of academic skills at issue in the IIMPaC program. For instance, the timed writing test does not provide a “real world” context for writing; neither does it provide opportunities for students to create drafts to discuss with teachers and other students prior to revising their work. To explore this possibility during their second year of research, the authors are devising new outcome measures, including a more diverse set of quality indicators and tests that allow opportunities for revision. They are also adapting the Writing Apprehension Test, which was designed for high school and college students, to make it more appropriate for middle school students. In addition, they are working to overcome procedural problems that led to unpreventable attrition in the current study, thus reducing the likelihood that modest, but significant, effects on students’ attitudes and performances would be measurable.

E. National Writing Project at Kent State University – *Analysis of the Effects on Teachers’ Practices and Student Performance of Teachers’ Participation in the Invitational Summer Institute*

This study compared instructional practices among teachers who either had or had not participated in the NOWP Invitational Summer Institute (SI) and performance of their students on each of three writing tasks, two of which were direct assessments of writing. The study also piloted an approach to identifying higher-order writing skills through a revision task in grades 5–12.

Teachers’ instructional practices were assessed through an analysis of teachers’ responses to the revision task. Student performance was assessed using holistic scoring based on measures derived from the writing section of the Ohio Graduation Test. Students were asked to identify which of the pieces in the portfolio they regarded as their best work, providing data on student assessments of quality writing.

The results revealed that teachers who had participated in the NOWP SI made comments regarding revision that were consistent with NOWP principles to a greater extent than did non-NOWP teachers. Specifically, in comparison with non-NOWP teachers, they focused more on audience than on mechanical aspects of writing and the tone of their comments was consistent with the idea that both students and teachers were members of a community of writers.

Analyses of student performance on the three assigned writing tasks revealed that students of NOWP-trained teachers obtained better scores on all three tasks than did students whose teachers had not participated in NOWP. However, as these results occurred for all three assessments—including the first assessment, which was meant to serve as a baseline measure—the superior performance of students of NOWP-trained teachers cannot be reliably attributed to their teachers’ participation in the program. Their better performance may be the result of a selection artifact, or it may be the result of the timing of the assessments.

Timing of the assessments may have had an effect because, for younger students, baseline assessments were conducted after several months of instruction. Thus, the higher scores of students of NOWP teachers on the initial assessment may be a result of differences in instruction that occurred before the baseline measure was taken. For older students in semester-long high school classes, however, the baseline measure was collected at the beginning of the course, making it a better basis for comparing the performance of students of teachers who either had or had not participated in the NOWP SI. In these classes, the performance of students of NOWP-trained teachers was similar to those of teachers who had not participated in NOWP at baseline, but the students of NOWP-trained teachers obtained better scores on the final assessment than did students of teachers who had not participated in NOWP. This finding indicates that, in circumstances where a better baseline measure was available, teachers’ participation in NOWP did appear to have a positive effect on student performance, but the small sample size made this finding inconclusive.

In addition, the study reports development of a revision task as an assessment instrument. Results on this task are inconclusive; careful analyses of student performance revealed considerable heterogeneity in responses, which may have overwhelmed potential group differences. This outcome points to the need for more fine-grained analyses of revision, which, given the emphasis on revision across multiple drafts in NOWP, may be more likely to reveal effects of the program.

V. LESSONS LEARNED – ANALYSIS OF LIMITATIONS AND BENEFITS OF THE LSRI APPROACH

A. Analysis of Methodological and Programmatic Limitations

In this section, we take a step back from the specific LSRI studies to highlight possible sources of the uneven results found in this first set of local site studies. We look at both methodological and programmatic issues and examine the implications of each for our work in the currently active second cohort of local site studies. These “lessons learned” offer insights into the kinds of challenges inherent in this approach to evaluating the effects of a complex, large-scale professional development endeavor.

LSRI was designed to take into account local variation in programming. LSRI was able to examine potential differences in the nature and degree of program effects by supporting writing project sites to frame the salient questions, articulate the expected

effect on student learning, and identify the most appropriate districts, schools, or classrooms to study. However, such a design approach also highlights two new challenges—tapping and building local expertise to conduct and learn from this research and providing appropriate technical assistance.

- **Timing.** In the first year, funding for LSRI closely followed the writing project’s overall funding cycle. Sites submitted proposals in late spring and were notified of funding in late summer. This schedule meant that sites had only a limited time to carry out critical preparatory work, such as finalizing research instrumentation, securing approval from university institutional review boards and school district research offices, and selecting program and comparison teachers to participate in the research. In two instances, this meant that pretest measures were not collected until November and December; therefore any student growth in writing which may have taken place during the first three months of the school year was not measured or subsumed within the comparisons in those sites. Perhaps more important, because baseline measures were collected well into the school year, they may have reflected the effects of instruction that occurred during that period, minimizing the likelihood of observing differences in student performance between the beginning and end of the observation period.

In the second LSRI cohort, we were able to notify funded sites much earlier (by the beginning of June) so that sites could complete pre-work and administer baseline writing prompts at the beginning of the school year.

- **Forming Appropriate Comparison Groups.** Identifying, recruiting, and maintaining appropriate comparison groups proved to be quite challenging for a variety of reasons. These include issues such as identifying a pristine comparison group either because of saturation of writing project programming within a site’s service area or because of spread of practice within a school, establishing a comparison group that both serves similar students and exhibits similar levels of professional expertise and motivation for pursuing professional growth, and securing the ongoing cooperation of teachers with the study.

Further, although teachers in comparison groups had not experienced the amount or intensity of professional development related to writing as those who participated in the programs offered by LSRI sites, they had, in several cases, participated in some form of professional development related to writing, making differences between program and comparison groups difficult to detect. Although related to the problem of finding “pristine” comparison groups, this issue is somewhat different in that, rather than dealing with concern about the diffusion of NWP practices, it deals with the potential for other programs to affect teacher practice and, as a result, student performance. As these alternative programs varied in content and level of teacher exposure both within and across sites, they have the potential to reduce the observed differences between program and comparison schools without offering a sound basis for assessing the results of participation in programs guided by NWP principles

relative to the potential effects of participation in other professional development programs.

In fact, we know of no professional development program that offers the systematic, intensive approach to the teaching of writing offered by NWP. Thus, even with increased attention to questions of the effects of NWP programs compared to the effects of other programs, such comparisons are unlikely to provide conclusive results. Assessing the comparative effects of NWP programs is likely to remain a matter of comparing those programs to “everything else,” where “everything else” constitutes a heterogeneous set of programs in which the level of exposure varies significantly within comparison schools and districts.

In the second LSRI cohort, we have addressed these challenges by assisting sites in prioritizing the characteristics of potential comparison groups; meeting with district administrators in order to assist sites in securing the administrators’ support for participation in local-site studies; and recommending that sites increase overall sample sizes and oversample comparison teachers to help address problems with attrition.

- **Appropriateness and Sensitivity of Available Measures of Student Writing.** One principle of this initiative is to employ direct assessments of student writing. Most widely available and tested measures of student performance on first, or, at least, early drafts of student papers. While many writing projects see this as one of many important indicators of skill and ability in writing, they have also identified a range of challenges with such measures. First, many of the scoring schemes use a relatively small number of scale points. While this approach increases the reliability of the assessment (and reduces its overall cost), it also means that there is less discrimination, and therefore, it is more difficult to measure growth. Second, measuring a single piece of first-draft writing makes it difficult to measure a range of valued skills such as revision and writing in a range of genres and for a variety of audiences.

To address these issues, the NWP has invested in developing an archive of writing prompts and rubrics that make accessible field-tested assessment schemes that can provide more discrimination and may therefore be better situated to measure growth over the course of the year.

- **Clear Articulation and Theorization of the Relationship Among Professional Development, Anticipated Change in Teaching Practices, and Expected Changes in Student Performance.** Writing project sites have a relatively long history of examining the ways in which professional development makes a difference for teachers’ knowledge, skills, and attitudes, and to some extent, even their practice. Systematic examination of student outcomes derived from changes in those teaching practices has not been part of the typical purview of research conducted by writing project sites. In many of the first cohort sites, the relationship between the teaching practices advocated and adopted and the type and amount of impact on student

learning was under-theorized. Without a clearly articulated theory about the connections between teacher participation in professional development programs oriented around NWP principles and student performance, it is difficult to interpret why a particular program seemed to have a positive impact or why there seemed to be no effect. Further, these unarticulated theories make it difficult to consider the full range of variables that may support or impede students' development as writers. This problem affected the research by supporting the examination of programmatic expressions that did not fully explicate or address the necessary conditions to ensure the effects sought.

The current group of LSRI studies is more sophisticated in the design and rigorous in implementation. The second round of studies successfully collected writing samples at the beginning of the relevant intervention; the studies are including more teachers and students in the sample; they are collecting a wider range of and deeper data about the classroom practices of both program and comparison teachers; and a few are adding additional measures of student outcomes.

B. Analysis of Benefits of the LSRI Approach

1. LSRI provided a significant opportunity to build writing project leaders' capacity to conduct research about the impact of their work on student outcomes.

As previously discussed, LSRI highlights a new challenge—tapping and building local expertise to conduct and learn from research. One significant accomplishment of this first year has been to build participating sites' capacity to design and conduct research that examines the effect of writing project work on teacher practices and the effects of this work on student learning. Through participating in the *NWP Research Forum* meetings and collaborating with senior NWP research staff, sites built their knowledge of designing quasi-experimental research studies, forming appropriate comparison groups, structuring writing assessments in ways that yield valid and reliable results, and developing and refining instruments to examine effects on classroom practice. In addition, sites are building relationships with faculty and other researchers in their universities who have specific expertise in research design, student assessment, and quantitative analysis.

Building writing project sites' capacity for conducting research of this type matters. Sites around the country are expressing interest in demonstrating the effect of their work on students' writing performance. Increasingly, sites are being challenged to provide research-based evidence that the services they provide make a difference. This includes the examination of the merit of the curricular and instructional ideas provided, as well as the efficacy of the means of fostering them. The response of the first cohort of sites to their results illustrates that they are interested in using them to examine the efficacy of their work. Beyond that, they are exploring ways in which the results can be used strategically to advocate best practice in instruction.

During the first year of the initiative, NWP found that expertise in designing and conducting experimentally based research studies was unevenly distributed across NWP sites. Therefore the NWP needed to invoke multiple means of supporting site-based research efforts. From the beginning, this research initiative was designed to provide technical assistance from the national office in the form of national meetings—the *NWP Research Forum*, where sites could learn from each other—and through site visits from senior NWP research staff. NWP staff learned that, in many places, the need for such support exceeded what was originally envisioned. In this regard, the NWP is no different from most other reform efforts facing the challenges of expanding its programs to include multiple sites that differ in their needs and capacities. Insofar as the vision of large-scale implementation embraces the ongoing evaluation and improvement of efforts as described above, the capacity to enact such a vision will remain a challenge, as has been noted in systematic analyses of efforts to scale up education reforms across a variety of programs that differ in their topical focus, the grade levels the programs are intended to address, and the approach to scale-up (e.g., highly controlled programs that employ similar curricular materials and methods of evaluation vs. programs based on more general principles and permit more adaptation to local circumstances). However, the NWP infrastructure and approach to mutual support within it offers a means of exploring unique approaches to pursuing this vision (Glennan et al. 2004).

2. LSRI provided the opportunity to identify and collect high quality research instruments for assessing student writing and measures of classroom practice.

Selecting high-quality research instrumentation for assessing student writing and classroom practice has proven to be a significant challenge. Therefore, the NWP is developing three related databases of materials that writing project sites can draw on—writing prompts that have been field-tested in large-scale assessment systems, rubrics for assessing student writing, and a range of instruments for measuring classroom practice (e.g., student and teacher surveys, observation protocols, interview schedules). Items are reviewed by experienced teachers of writing, curriculum experts, and researchers and recommended for inclusion in the archive for general writing project use or for research.

3. LSRI pinpointed areas where national resources can be used to support local writing project sites' research.

- **National scoring of student writing.** In the first cohort, sites allocated much of their resources to scoring student writing. The cost of training and intellectual challenges of achieving high-quality scoring, as well as credibility issues for local scoring, even when very carefully carried out led us to establish a national scoring process whereby funded sites can submit student writing samples for common scoring. In this way, we can assemble significant national expertise to select benchmark papers and train scorers. Sites will not have to score their own students' writing leading to greater credibility of results. The arrangement provides for higher-quality assessment of student performance while avoiding the duplication required by developing assessments in each of the several sites. Participation in the national scoring activity is extremely powerful professional

development in its own right, greatly benefiting the participants as individuals while enhancing the capacities of the sites in which they work. Finally, this approach further builds sites' ability to conduct high-quality research by freeing up local resources to carry out other aspects of their research designs.

- **More extensive individualized site consulting.** NWP senior research staff members are also providing more extensive site-based technical assistance during all phases of the research process. In the first year, following initial conference calls and email exchanges, sites typically scheduled a single visit, most of which took place well into the school year. In contrast, research staff and second cohort sites have already completed at least one visit just prior to or at the beginning of the school year to finalize start-up issues such as recruitment of program and comparison groups, procedures for administering pretest writing samples, and design of classroom practices measures. In addition, mid-year visits are scheduled to address issues such as ongoing qualitative data analysis and the establishment of data bases for recording quantitative data. This type of support will assist sites in analyzing and reporting their results in the months to come.
- **Building stronger relationships between site leadership and local research talent.** In reviewing second cohort proposals, we gave priority to sites that set aside resources for working with local research partners with expertise in quantitative analysis, experimental design, or student assessment. Our work with site leadership to expand their understanding of experimental design and analysis of student assessment data has deepened their capacity to work effectively with local research partners in shaping designs that are consistent with local programming goals and approaches and in interpreting and learning from results. Over the long run, we anticipate that sites will be able to build on these relationships to conduct ongoing research about the effectiveness and impact of their work.

VI. FUTURE DIRECTIONS – DESIGNING A NATIONAL STUDY

The second cohort of LSRI studies is currently underway and, as noted above, the study designs have benefited from the lessons learned in Cohort I. As we look across these five studies from Cohort I, we note modest positive results and recognize various methodological and programming limitations. The LSRI is building capacity in writing project sites to evaluate program effects and to incorporate research-based learning into ongoing program development and implementation.

We are confident that the full collection of studies in Cohort I and II will be very useful in planning the focus and design of a new national study. We seek to learn more from the entire set of LSRI studies to help identify an appropriate focus for a national study, articulate the local dimensions and variations of writing project practices that could be expected to lead to changes in teacher practices and student learning, and use local leverage to secure appropriate comparative reference. Building on this local capacity could further support the rigor of a national study.

VII. REFERENCES

- Culham, R. 2003. *6 + 1 Traits of Writing*. New York, NY: Scholastic Professional Books.
- Daly, J. A., and M. D. Miller. 1975. "The Empirical Development of an Instrument to Measure Writing Apprehension." *Research in the Teaching of English* 9, 242–249.
- Glennan, T. K. Jr., S. J. Bodilly, J. R. Galegher, and K. A. Kerr. 2004. *Expanding the Reach of Education Reforms*. Santa Monica, CA: RAND Corporation.
- National Research Council. 2002. *Scientific Research in Education*. Washington, D.C.: National Academy Press.