

NATIONAL WRITING PROJECT

LOCAL SITE RESEARCH INITIATIVE REPORT

Cohort II

2004–2005

Part I – Summary Report of National Results

JANUARY 2006

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. Margaret Perrow served as consulting research editor and, in that capacity, edited each research report and prepared the written summaries that form a major portion of this report.

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

Gateway Writing Project, University of Missouri, St. Louis, MO

Principal investigators: Nancy Robb Singer and Diane Scollay

Research team: Astra Cherry, Michael Lowenstein, Tom Cornell, Jon Marshall

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

Principal investigator: Sherry Seale Swain, National Writing Project

Research team: David Morse, Richard L. Graves, Linda Irby, Marty Clark, Dolyene Davis, Mandy Goldman, Patricia Parrish, Liz Townsend, Spike Harris

New York City Writing Project, Lehman College, The City University of New York, Bronx, NY

Principal investigators: Marcie Wolfe, Anne Campos, and Nancy Mintz

Research team: Roger Peach, Joe Bellacero, Debra Freeman, Virginia Moss, Ed Osterman, Patsy Wooters

Southern Nevada Writing Project, University of Nevada, Las Vegas, NV

Principal investigators: Marilyn McKinney and Saralyn Lasley

Research team: Michael Nussbaum, Rosemary Holmes-Gull, Shawn Kelly, Kim Sicurella, Margaret Thackeray, and Mary Kinghorn

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

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

Research team: Eddi Christensen, Tim Dewar, Sarah Hochstetler, Henny Kim, Chia-Chen Lai, Suzie Null, Nada Rayyes, Paul Rogers, Jennifer Scalzo

A panel of distinguished scholars with expertise in research served as external reviewers of the work: Anne DiPardo of the University of Iowa; JoAnne T. Eresh, of ACHIEVE, Inc.; Sandra Murphy, of the University of California, Davis; Gail Offen-Brown, of the University Writing

Center, University of California, Berkeley; Barbara Storms, of California State University, East Bay; and Melanie Sperling of the University of California, Riverside.

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
January 2006

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	OVERVIEW OF NWP LOCAL SITE RESEARCH INITIATIVE	
	A. Proposal Selection Process.....	1
	B. New Features of LSRI in 2004–2005.....	2
III.	SUMMARY OF LSRI RESEARCH DESIGN AND METHODS	
	A. Program Emphasis	3
	B. Grade Levels Examined	4
	C. Research Approaches	4
	D. Framework for Evaluating and Scoring Student Writing.....	4
IV.	SUMMARY OF RESULTS	
	A. Overall Results for Student Writing Performance	6
	Table 1: Summary of Student Writing Performance Across Sites.....	6
	B. Summary of Results by Site	7
	Table 2: Summary of Results by Site	8
	1. Gateway Writing Project, University of Missouri, St. Louis	13
	2. Mississippi Writing/Thinking Institute, Mississippi State University.....	14
	3. New York City Writing Project, Lehman College, City University of New York	15
	4. South Coast Writing Project, University of California, Santa Barbara	16
	5. Southern Nevada Writing Project, University of Nevada, Las Vegas.....	17
V.	LESSONS LEARNED	
	A. LSRI Design Changes from Year One to Year Two.....	19
	B. Research Design Challenges	19
	C. Opportunity to Add Expertise and Knowledge to the Field.....	20
VI.	CONCLUSION	21
VII.	REFERENCES	23
VIII.	APPENDIX A: Inter-rater Reliabilities by Trait and by Level	24

NATIONAL WRITING PROJECT

JANUARY 2006

LOCAL SITE RESEARCH INITIATIVE REPORT – Cohort II (2004–2005)

I. INTRODUCTION

This report describes the activities and results of the second year of the Local Site Research Initiative (LSRI) of the National Writing Project (NWP). NWP structured the LSRI as a national portfolio of locally defined research efforts. This research was designed to study the effectiveness of the writing project model and its impact on students in a range of contexts. Four studies subsumed within the LSRI portfolio focused on the effects of intensive professional development partnerships between writing project sites and schools or school districts. Specifically, these studies focused on NWP outreach to teachers in professional development settings. The fifth study examined the effects of teachers' continued involvement with the writing project following the invitational summer institute. The five studies were conducted by five writing project sites in diverse settings: Gateway Writing Project (University of Missouri, St. Louis), Mississippi Writing/Thinking Institute (Mississippi State University), New York City Writing Project (Lehman College), South Coast Writing Project (University of California, Santa Barbara), and Southern Nevada Writing Project (University of Nevada, Las Vegas). All five studies employed quasi-experimental, comparative designs. In this report, we share the results of these studies and identify important learnings that inform the National Writing Project's plans for national evaluation and its ongoing research and evaluation support to local sites.

II. OVERVIEW OF NWP LOCAL SITE RESEARCH INITIATIVE

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 varies according to local needs, reform priorities, school conditions, and research contexts. Each writing project site emphasizes common core principles of effective instruction, while the design and delivery of services is negotiated with local education authorities. NWP's goal for the LSRI is to develop a growing body of research that examines local professional development programs based on these core principles and that illuminates teacher practices and student achievement in writing across a range of grade levels, schools, and local contexts.

A. Proposal Selection Process

National Writing Project sites were invited to participate through a request for proposals, and those selected were each awarded \$20,000 and technical assistance from the national office. Twenty writing project sites submitted proposals for research to be conducted during the 2004–2005 academic year. All proposals were reviewed both by NWP staff and by distinguished scholar researchers serving as external 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 funding were chosen both for the quality of their research plans and for their ability to contribute to a collective portfolio whose various pieces complement each other with respect to the questions asked, the contexts in which those questions are explored, and the methods used to explore them. At the core, each site's study was designed to investigate how key components of its program activities contributed to changes in teaching practices as well as to student learning and achievement in writing. The studies employed various complementary methods of data collection and analysis to examine student learning as represented in performance on state and local assessments, responses to writing prompts and tasks, and portfolios.

Finally, NWP offered an extensive technical assistance program to further develop 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 that writing projects serve.

B. New Features of LSRI in 2004–2005

In the first set of LSRI research studies (National Writing Project 2005), NWP and participating sites encountered a number of methodological challenges. In order to address these issues in supporting this second set of studies, NWP added several additional support structures in 2004–2005:

- **Timing.** The application and notification process for NWP's second panel of LSRI sites took place considerably earlier, prior to the beginning of June. This earlier timetable allowed sites to complete research planning and preparation (e.g., development and selection of instruments, completion of Institutional Review Board processes, as well as recruitment of program and comparison groups) during the summer, and enabled sites to administer baseline measures as early as possible in the school year.
- **Increased access to field-tested measures of student writing performance and classroom practice.** As documented in the year 1 report, the first LSRI sites individually faced the daunting task of developing their own measures of student writing performance and of classroom practice. In order to have data of sufficient technical quality to support this research, it was necessary to provide increased support for the assessment of student writing. The NWP created an archive of field-tested student writing prompts and frameworks for evaluating student writing. This archive drew on major assessment systems in the United States, Canada, the United Kingdom, Australia, and New Zealand. Four of the five studies reported on here employed writing prompts from this archive. In

addition, the NWP began building an archive of measures of classroom practice (including surveys, inventories of practice, and interview and observation protocols). Several sites turned to these as models and sources of questions as they developed their own classroom-practice measures.

- **National scoring of student writing.** To ensure technical rigor and credibility of the writing assessment data, scoring and data processing were conducted nationally and independently of the local sites. All student writing samples were scored using a common evaluative framework, a modified version of the Six+1 Trait Writing Model (Bellamy 2005). A national panel of experts in the assessment of student writing, along with NWP senior researchers, made the framework more appropriate for research use by extending the scale to six points, clarifying the language throughout, and focusing evaluative judgments exclusively upon the written product. All student writing samples, with identifying information masked by a coding system, were then scored at a national scoring conference. (Technical details of the assessment system and attendant scoring are presented in section III below.)

III. SUMMARY OF LSRI RESEARCH DESIGN AND METHODS

Four of the five studies included in this second panel analyzed the effect of professional development inservice programs offered by writing project sites to schools and districts, and one study analyzed the effect of teachers' continuing involvement in writing project programming. 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 kinds and differing levels of involvement, and ranged from an investigation of the impact of participation in the first year of a professional development program to a summary study of the impact of a multiyear district partnership.

A. Program Emphasis

Four studies—those of the Gateway Writing Project, the Mississippi Writing/Thinking Institute (MWTI), the New York City Writing Project (NYCWP) and the South Coast Writing Project (SCWriP)—examined inservice work offered by writing project teacher-consultants to schools or districts in their local service areas. The Gateway Writing Project examined the impact of an intensive leadership development program focused on the teaching of writing, offered with the support of a suburban St. Louis school district, which took place during the year of the study. MWTI examined the effects of an intensive professional development program for ninth grade English teachers that occurred in districts in which seventh and eighth grade teachers had received complementary training in previous years. NYCWP examined the influence of school-based inservice programming on the classroom practices of teachers across several disciplines and on students in English classrooms, with a special focus on English language learners. SCWriP examined the impact of inservice programming on English language arts teachers who were active participants in school-based programming for one to three years. The fifth site, the

Southern Nevada Writing Project, examined the effects of participation in its Family Writing Project, which engages students and their families in writing activities outside of school and provides teachers opportunities for continued involvement with the writing project following the summer institute.

B. Grade Levels Examined

Collectively, the studies covered a full range of grade levels. The Gateway study focused on elementary school students; SCWriP studied implementation at the upper elementary and middle school levels; and Southern Nevada’s study focused on middle school students. The MWTI and NYCWP studies were conducted in high school settings.

C. Research Approaches

Central to each of the five studies was the site’s commitment to understand what difference writing project professional development makes for participating teachers’ practices and, in turn, what difference those changes in instructional practices make for student learning. Each study, as required by the LSRI guidelines, employed direct assessments of student writing, and each included some form of comparative reference against carefully matched comparison classes and/or students.

All five studies employed a quasi-experimental design and analyzed, at minimum, pre/post writing samples. The MWTI study matched program and comparison schools. The Gateway, NYCWP, and SCWriP studies matched individual teachers and their classrooms. The SNWP study matched individual students who participated in the Family Writing Project with individual students who did not participate.

D. Framework for Evaluating and Scoring Student Writing

Building upon a long tradition of writing assessment, the NWP defined a rigorous evaluation framework: developed standards and related developed “anchor papers”—samples exemplifying each level of achievement—along with descriptive commentary; and designed and conducted an independent national scoring conference for scoring student written work. All five studies participated in the conference, submitting student writing samples—with identifying information removed—to be scored.

Rubric. The evaluative framework included a rubric attending to six specific attributes as well as the overall character of students’ writing, adapted from the widely used 6+1 Traits Writing Model (Culham 2003). The six specific attributes are as follows:

- Ideas / Content Development (Idea): establishing purpose, selecting and integrating ideas, including details to support, develop, or illustrate ideas
- Organization (Org): creating an opening and closing, maintaining focus, ordering and relating events, ideas, and details to provide coherence and unity in the writing
- Voice (Voice): communicating in an engaging and expressive manner, revealing the writer’s stance toward the subject
- Sentence Fluency (Sent Flu): constructing sentences to convey meaning, controlling syntax, and creating variety in sentence length and type
- Word Choice (Wrd Chc): choosing words and expressions for appropriateness, precision, and variety

- Conventions (Conv): controlling grammar, punctuation, spelling, capitalization, and paragraphing. (Bellamy 2005)

In addition to scores in each of these areas, each writing sample received an overall holistic score, one defined not as an aggregate of these component parts but as an independent, overall, summary judgment.

A national panel of experts on the assessment of student writing, along with NWP senior researchers, determined that the 6+1 Trait Model of Writing, while sufficiently comprehensive, required certain modifications to make it more appropriate for use in research studies. The following modifications were made in the rubric prior to the scoring conference:

- The scale of the rubric was extended from four to six points in order to ensure sufficient discrimination and therefore to allow increased sensitivity to any changes in student performance.
- The language defining the traits was clarified to enhance the reliability of evaluative judgments.
- The evaluative judgments were modified to focus exclusively upon the student writing (where, on occasion, the original rubric included references to the reader's reactions or to the writer's personality as the basis for judgment).

Scoring. The student writing samples were scored at a national scoring conference held on June 16 to 18, 2005. Students' writing samples from all five sites were integrated and scored—the local site research teams determined which set of grade level standards (elementary, middle, or high school) would be applied when scoring work from their sites. All of the student writing was coded, with identifying information removed so that scorers could not know any specifics of the writing sample being evaluated (e.g., site of origin, group [program or comparison], or time of administration [pretest or posttest]). Across all sites, 23% of the student writing was scored twice (so that reliabilities could be calculated): 10% of the high school papers, 52% of the middle school, and 11% of the elementary were scored twice.

The scorers participated in six hours of training at the beginning of the scoring conference. Their scoring was calibrated to a criterion level of performance at that time, and then recalibrated following every major break in the scoring (meals and overnight). Overall, reliabilities (measured as interrater agreement, defining agreement as two scores being identical or within one single score point of each other) ranged from 90% to 95%, with an aggregate across all scores of 92%. This level of reliability is as high as is typically observed, and easily adequate to support the research purposes pursued here. (See appendix A for complete data describing the reliability of the scoring of student writing.)

IV. SUMMARY OF RESULTS

A. Overall Results for Student Writing Performance

Table 1 summarizes findings about student writing performance across studies. On table 1, upward-pointing triangles (in green) denote positive findings favoring the National Writing Project. Downward pointing triangles signal findings favoring the comparison group. Solid triangles indicate statistically significant differences. Table 1 shows results for each site and includes findings comparing the NWP group to the comparison group for each of the six attributes of writing as well as the independent overall holistic score.

Table 1: Summary of Student Writing Performance Across Sites

	Ideas	Org	Voice	Sentence Fluency	Word Choice	Conventions	Holistic Score
Gateway	▲	▲	▲	▲	△	▲	▲
MWTI	▲	▲	▲	▲	▲	▲	▲
NYCWP	▲	▲	▲	△	▲	▲	▲
SCWriP	△	△	△	△	△	△	△
SNWP	△	△	△	△	▲	△	△

▲ - Favorable Results – Significant; △ - Favorable Results – Not Significant; ○ - No Difference; ▽ - Unfavorable Results – Not Significant; ▼ - Unfavorable Results – Significant

In every one of the comparisons (seven scores and five sites, 35 contrasts in all), the results favor the students in classes of teachers participating in the NWP. That is, for every one of the measured attributes of writing and in every site, the improvement of students taught by NWP participating teachers exceeded that of students in classes whose teachers were not participants. Moreover, in 20 of the 35 contrasts (57%) the differences between NWP participants’ students and comparison students were statistically significant. The results, taken across sites, indicate a consistent pattern favoring the NWP. (Although the studies employed a common evaluative framework and writing was scored under the same conditions, comparisons of sites to each other are discouraged because of substantial differences in sites’ context and overall research design, and differences in the writing prompts used and the conditions under which they were administered.)

Two additional patterns in these results are particularly noteworthy. The first is the consistently favorable results for the attributes of Ideas, Organization, and Voice. These command attention because the quality of ideas and their expression are particular foci of writing project efforts. These results signal effectiveness in the very areas that are emphasized in many writing project activities.

The second noteworthy pattern in these data is the consistently positive outcomes for Conventions (such as grammar, punctuation, or capitalization). It is inherent in writing project programming that conventions are best addressed in the context of ideas and their expression—not as an abstract pursuit in and of themselves. The positive findings for Conventions

corroborate the validity of this view. Nowhere does the pursuit of the quality of ideas redound to the detriment of basic skills such as grammar, punctuation, or capitalization.

B. Summary of Results by Site

This section summarizes each site's research design and results, first in table 2 and then in the form of narrative analytic summaries. Table 2 reports the results of each study, including the impact on students in NWP classes across sites. Summarized there are the program emphases, the research designs, and the results of each study. The summary of results in the right-hand-most column lists the changes in mean performance from pretest to posttest (on a six-point scale), first for the program group and then for the comparison group. While it can be readily seen that all comparisons favor the performance of the students in the NWP-supported programs, those that are statistically significant are identified by appropriate footnoting. Since all work was scored at once using the same evaluative framework, these differences are scaled similarly across sites.

The general pattern of results that favor the program group across sites is elaborated in table 2 below. Here we see that the scores of the program group tended to increase around one-half a scale point (over 8% of the full range of the scale and a much greater proportion of the available scale given initial performance) from start to end of the period of the research. By contrast, the comparison group students, even across contexts, tended to remain flat in their performance (though in a few instances even exhibiting lower scores at the end of the period).

Table 2: Summary of Results by Site

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

Program Emphasis	Research Design			Results																								
	Outcome Studied	Measures	Samples																									
Professional development partnership with 4 elementary and middle schools that serve high-poverty populations with a substantial proportion of English language learners	<p><i>Student Outcomes</i></p> <ul style="list-style-type: none"> • Scores on prompted writing exercise administered at 2 points in time • Student attitudes toward writing <p><i>Teacher Outcomes</i></p> <ul style="list-style-type: none"> • Teacher classroom practices 	<p><i>Student Outcome Measures</i></p> <ul style="list-style-type: none"> • Writing prompts drawn from an archive that includes retired tasks from established writing assessments and independently scored using the 6-trait scoring rubric at a national scoring conference <p>• Writing Apprehension Test (Daly & Miller 1975)</p> <p><i>Teacher Outcome Measures</i></p> <ul style="list-style-type: none"> • Teacher evaluations of the professional development (interview protocol) • Semistructured interviews and classroom observations to determine effects on teachers' classroom practices • Collections of teacher and student work to illustrate approaches to classroom practices 	<p>Students in grades 4 to 8 and their English language arts teachers</p> <p><i>Program</i> Students = 149 Teachers = 8</p> <p><i>Comparison</i> Students = 74 Teachers = 7</p>	<p style="text-align: center;">Mean change from premeasure to postmeasure ¹</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>Ideas</th> <th>Org</th> <th>Voice</th> <th>Sent Flu</th> <th>Wrđ Chc</th> <th>Conv</th> <th>Holistic</th> </tr> </thead> <tbody> <tr> <td>Prog.</td> <td>.31</td> <td>.26</td> <td>.18</td> <td>.21</td> <td>.15</td> <td>.08</td> <td>.23</td> </tr> <tr> <td>Comp.</td> <td>.11</td> <td>.09</td> <td>.09</td> <td>.01</td> <td>-.07</td> <td>-.03</td> <td>.10</td> </tr> </tbody> </table> <p>¹ Cell entries represent mean differences (postmeasure score minus premeasure score) on a 6-point scale over the term of the program.</p> <ul style="list-style-type: none"> • The program students' writing scores increased more than comparison students' scores on the holistic and all 6 analytic measures. These increases were not large enough to be statistically significant. • Although both program and comparison teachers taught students strategies for preparing to write, program students actually increased their use of these strategies, such as marking up the text to which they were responding and prewriting, to a statistically significant degree. • Program teachers taught students to develop ideas and revise the content of their writing as well as to edit and proofread it. In contrast, comparison teachers focused primarily on teaching students to proofread their writing. • Program teachers used a number of practices to promote students' investment in their own writing, including providing opportunities for students to select their own topics, collect all of their writing in portfolios, and discuss their writing with other students. 		Ideas	Org	Voice	Sent Flu	Wrđ Chc	Conv	Holistic	Prog.	.31	.26	.18	.21	.15	.08	.23	Comp.	.11	.09	.09	.01	-.07	-.03	.10
	Ideas	Org	Voice	Sent Flu	Wrđ Chc	Conv	Holistic																					
Prog.	.31	.26	.18	.21	.15	.08	.23																					
Comp.	.11	.09	.09	.01	-.07	-.03	.10																					

MO – Gateway Writing Project, University of Missouri, St. Louis

Program Emphasis	Research Design			Results																								
	Outcome Studied	Measures	Samples																									
Professional development partnership with midsize school district, focus on leadership development re teaching of writing	<p><i>Student Outcomes</i></p> <ul style="list-style-type: none"> Standardized English language arts test scores Scores on prompted writing exercise administered at 4 points in time <p><i>Teacher Outcomes</i></p> <ul style="list-style-type: none"> Teachers’ classroom practices Teacher knowledge of best practice Teachers’ attitudes towards writing and writing instruction 	<p><i>Student Outcome Measures</i></p> <ul style="list-style-type: none"> Reading assessment scores, Gates-MacGinitie Writing prompts drawn from the district’s writing assessment program and independently scored using a 6-trait scoring rubric and a holistic rubric at a national scoring conference <p><i>Teacher Outcome Measures</i></p> <ul style="list-style-type: none"> Assessment of teacher knowledge of best practice Surveys and interviews on teacher attitudes toward writing and writing instruction and on teacher instructional practice Observations of teachers’ writing instruction 	<p>Students in grades 3 to 5 and their classroom teachers</p> <p><i>Program</i> Students = 81 Teachers = 7</p> <p><i>Comparison</i> Students = 67 Teachers = 7</p>	<p style="text-align: center;">Mean change from premeasure to postmeasure¹</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>Ideas²</th> <th>Org³</th> <th>Voice²</th> <th>Sent Flu²</th> <th>Wrđ Chc</th> <th>Conv³</th> <th>Holistic²</th> </tr> </thead> <tbody> <tr> <td>Prog.</td> <td>.42</td> <td>.54</td> <td>.50</td> <td>.59</td> <td>.46</td> <td>.61</td> <td>.48</td> </tr> <tr> <td>Comp.</td> <td>-.07</td> <td>-.01</td> <td>-.06</td> <td>-.18</td> <td>.15</td> <td>-.10</td> <td>-.03</td> </tr> </tbody> </table> <p>¹ Cell entries represent mean differences (postmeasure score minus premeasure score) on a 6-point scale over the term of the program. ² Difference in the change between groups over time is significant at $\alpha < .05$. ³ Difference in the change between groups over time is significant at $\alpha < .01$.</p> <ul style="list-style-type: none"> The program students’ writing scores increased more than comparison students’ scores on the holistic and all 6 analytic measures. All differences, except for word choice, are statistically significant. Differences for organization and conventions are statistically significant at $\alpha \leq .01$ while differences in holistic scores, Ideas, Voice, and Sentence Fluency are significant at $\alpha \leq .05$. Program students’ scores on the Gates-MacGinitie reading test also increased more than comparison students’ scores. These differences were statistically significant. Teachers in both groups emphasized the development of students’ vocabulary and made literacy resources readily accessible in their classroom. Program teachers were observed engaging their students in a wider range of writing tasks, designing writing instruction that extended the development of a piece of writing over multiple class periods (sometimes lasting for weeks or months), and modeling explicit connections between reading and writing. In contrast, writing instruction observed in comparison teachers’ classrooms largely centered on completing tasks that mirrored the types of timed writing on state tests. 		Ideas ²	Org ³	Voice ²	Sent Flu ²	Wrđ Chc	Conv ³	Holistic ²	Prog.	.42	.54	.50	.59	.46	.61	.48	Comp.	-.07	-.01	-.06	-.18	.15	-.10	-.03
	Ideas ²	Org ³	Voice ²	Sent Flu ²	Wrđ Chc	Conv ³	Holistic ²																					
Prog.	.42	.54	.50	.59	.46	.61	.48																					
Comp.	-.07	-.01	-.06	-.18	.15	-.10	-.03																					

MS—Mississippi Writing/Thinking Institute (MWTI), Mississippi State University

Program Emphasis	Research Design			Results																								
	Outcome Studied	Measures	Samples																									
Professional development partnership with 2 schools, content focused on writing improvement program	<p><i>Student Outcomes</i></p> <ul style="list-style-type: none"> Scores on pre/post on-demand writing exercises <p><i>Teacher Outcomes</i></p> <ul style="list-style-type: none"> Teachers' implementation of individual program-sponsored practices as well as a holistic assessment of their overall degree of implementation 	<p><i>Student Outcome Measures</i></p> <ul style="list-style-type: none"> Two writing prompts, one drawn from an archive that includes retired tasks from established writing assessments, the second designed to be equivalent to the first. Writing was independently scored using the 6-trait scoring rubric at a national scoring conference <p><i>Teacher Outcome Measures</i></p> <ul style="list-style-type: none"> Teacher interviews and classroom audits rated and analyzed to measure implementation of MWTI practices. Description of MWTI professional development in each school for each succeeding year along with data on teacher participation and turn-over. 	<p>9th grade students and their English teachers in program and comparison schools</p> <p><i>Program Schools</i> Students = 298 Teachers = 5</p> <p><i>Comparison Schools</i> Students = 157 Teachers = 5</p>	<p style="text-align: center;">Mean change from premeasure to postmeasure ¹</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>Ideas²</th> <th>Org₂</th> <th>Voice₂</th> <th>Sent Flu²</th> <th>Wrđ Chc²</th> <th>Conv²</th> <th>Holistic²</th> </tr> </thead> <tbody> <tr> <td>Prog.</td> <td>0.5</td> <td>0.6</td> <td>0.6</td> <td>0.3</td> <td>0.4</td> <td>0.4</td> <td>0.5</td> </tr> <tr> <td>Comp.</td> <td>0.0</td> <td>0.0</td> <td>-0.1</td> <td>-0.1</td> <td>-0.1</td> <td>0.1</td> <td>-0.1</td> </tr> </tbody> </table> <p>¹ Cell entries represent mean differences (postmeasure score minus premeasure score) on a 6-point scale over the term of the program. ² Difference in the change between groups over time is significant at $\alpha < .001$.</p> <ul style="list-style-type: none"> The program students' writing scores increased more than comparison students on holistic and all 6 analytic measures. All differences are statistically significant at $\alpha \leq .001$. Program teachers reported implementing teaching strategies presented during MWTI professional development. Analysis revealed that these teaching practices were more prevalent in program classrooms than in comparison classrooms. Statistically significant differences were found in favor of the program teachers in the areas of student choice and peer response. 		Ideas ²	Org ₂	Voice ₂	Sent Flu ²	Wrđ Chc ²	Conv ²	Holistic ²	Prog.	0.5	0.6	0.6	0.3	0.4	0.4	0.5	Comp.	0.0	0.0	-0.1	-0.1	-0.1	0.1	-0.1
	Ideas ²	Org ₂	Voice ₂	Sent Flu ²	Wrđ Chc ²	Conv ²	Holistic ²																					
Prog.	0.5	0.6	0.6	0.3	0.4	0.4	0.5																					
Comp.	0.0	0.0	-0.1	-0.1	-0.1	0.1	-0.1																					

NV – Southern Nevada Writing Project, University of Nevada, Las Vegas

Program Emphasis	Research Design			Results																								
	Outcome Studied	Measures	Samples																									
Family Writing Project, which provides direct services to students and their families and provides ongoing professional development to teachers who have already participated in the SNWP Invitational Institute	<p><i>Student Outcomes</i></p> <ul style="list-style-type: none"> • Student writing performance, using pre/post assessments of writing • Student attitudes toward writing <p><i>Teacher Outcomes</i></p> <ul style="list-style-type: none"> • Changes in instructional practice • Relationships between teachers and parents/families 	<p><i>Student Outcome Measures</i></p> <ul style="list-style-type: none"> • Writing prompts drawn from an archive that includes retired tasks from established writing assessments and independently scored using the 6-trait scoring rubric at a national scoring conference <p><i>Teacher Outcome Measures</i></p> <ul style="list-style-type: none"> • Student surveys and interviews of attitudes toward school and writing • Teacher interviews regarding classroom practices and teacher/family relations • Student survey responses regarding classroom practices 	<p>Students in grades 6 to 8, 4 teachers who facilitate the Family Writing Project</p> <p><i>Program</i> Students = 21 Teachers = 4</p> <p><i>Comparison</i> Students = 21</p>	<p style="text-align: center;">Mean change from premeasure to postmeasure ¹</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>Ideas</th> <th>Org</th> <th>Voice</th> <th>Sent Flu</th> <th>Wrđ Che²</th> <th>Conv</th> <th>Holistic</th> </tr> </thead> <tbody> <tr> <td>Prog.</td> <td>.21</td> <td>.38</td> <td>.50</td> <td>.24</td> <td>.50</td> <td>.41</td> <td>.28</td> </tr> <tr> <td>Comp.</td> <td>.12</td> <td>.15</td> <td>.07</td> <td>.16</td> <td>.00</td> <td>.03</td> <td>.09</td> </tr> </tbody> </table> <p>¹ Cell entries represent mean differences (postmeasure score minus premeasure score) on a 6-point scale over the term of the program. ² Difference in the change between groups over time is significant at a < .05.</p> <ul style="list-style-type: none"> • The program students’ writing scores increased more than comparison students on the holistic and all 6 analytic measures. The differences for word choice were statistically significant at $\alpha \leq .05$. • Program students’ positive attitudes toward writing held steady or increased while those of comparison students’ declined. Program students also reported writing more frequently. • Teachers involved in the Family Writing Project reported deepening their understanding of how writers write, improving their relationships with their students and students’ families, and developing their leadership and teaching skills. 		Ideas	Org	Voice	Sent Flu	Wrđ Che ²	Conv	Holistic	Prog.	.21	.38	.50	.24	.50	.41	.28	Comp.	.12	.15	.07	.16	.00	.03	.09
	Ideas	Org	Voice	Sent Flu	Wrđ Che ²	Conv	Holistic																					
Prog.	.21	.38	.50	.24	.50	.41	.28																					
Comp.	.12	.15	.07	.16	.00	.03	.09																					

NY – New York City Writing Project, Lehman College

Program Emphasis	Research Design			Results																								
	Outcome Studied	Measures	Samples																									
Professional development partnerships with 6 high schools in which NYCWP has worked from 1 to 5 years	<p><i>Student Outcomes</i></p> <ul style="list-style-type: none"> Scores on pre/post assessments of student writing <p><i>Teacher Outcomes</i></p> <ul style="list-style-type: none"> Teachers' knowledge/implementation of best practice 	<p><i>Student Outcome Measures</i></p> <ul style="list-style-type: none"> Writing prompts drawn from an archive that includes retired tasks from established writing assessments and independently scored using the 6-trait scoring rubric at a national scoring conference <p><i>Teacher Outcome Measures</i></p> <ul style="list-style-type: none"> Interviews/surveys of teachers' instructional practice and attitudes 	<p>Students in grades 9 to 12</p> <p><i>Program</i> Students = 42 (16 of whom are ELL students) Teachers = 8</p> <p><i>Comparison</i> Students = 39 (18 of whom are ELL students) Teachers = 3</p>	<p style="text-align: center;">Mean change from premeasure to postmeasure ¹</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th></th> <th>Ideas²</th> <th>Org²</th> <th>Voice²</th> <th>Sent Flu</th> <th>Wrđ Chc²</th> <th>Conv²</th> <th>Holistic²</th> </tr> </thead> <tbody> <tr> <td>Prog.</td> <td>0.5</td> <td>0.4</td> <td>0.8</td> <td>0.3</td> <td>0.2</td> <td>0.4</td> <td>0.6</td> </tr> <tr> <td>Comp.</td> <td>-0.3</td> <td>-0.3</td> <td>-0.4</td> <td>0.0</td> <td>-0.4</td> <td>-0.2</td> <td>-0.1</td> </tr> </tbody> </table> <p>¹ Cell entries represent mean differences (postmeasure score minus premeasure score) on a 6-point scale over the term of the program. ² Difference in the change between groups over time is significant at $\alpha < .05$.</p> <ul style="list-style-type: none"> The program students' writing scores increased more than comparison students on the holistic and all 6 analytic measures. All differences, except for sentence fluency, are statistically significant $\alpha \leq .01$. Subgroup analyses for ELL students demonstrated a similar pattern with program students' writing scores increasing more than comparison students' on all measures. These differences, except for sentence fluency and word choice, were statistically significant at $\alpha \leq .05$. Program teachers reported that they had adopted views and practices consistent with NYCWP professional development; in particular they treat writing as a process, use free writing, and engage students in responding to each others' writing. In addition, teachers in content areas other than English are incorporating writing into their instructional practice. 		Ideas ²	Org ²	Voice ²	Sent Flu	Wrđ Chc ²	Conv ²	Holistic ²	Prog.	0.5	0.4	0.8	0.3	0.2	0.4	0.6	Comp.	-0.3	-0.3	-0.4	0.0	-0.4	-0.2	-0.1
	Ideas ²	Org ²	Voice ²	Sent Flu	Wrđ Chc ²	Conv ²	Holistic ²																					
Prog.	0.5	0.4	0.8	0.3	0.2	0.4	0.6																					
Comp.	-0.3	-0.3	-0.4	0.0	-0.4	-0.2	-0.1																					

1. Gateway Writing Project, University of Missouri, St. Louis:
Leadership in Writing Institute: An Evaluation of Professional Development Impact

This study by the Gateway Writing Project at the University of Missouri-St. Louis investigated a yearlong professional development program implemented in partnership with a midsized school district. The program explicitly encouraged participating teachers to take an inquiry stance toward writing pedagogy by collecting and analyzing data from their own classrooms, reflecting on their practice, as well as trying and examining new classroom strategies. This study examined the professional development program's effect on teacher practices, and consequently on student achievement in writing.

The study focused on seven teachers of third through fifth grade and the students in their classes. Seven matched teachers and their students, drawn from the same district, provided comparison data. Teacher outcomes included knowledge of writing-process pedagogy, classroom practices, and attitudes toward writing and writing instruction. These outcomes were measured by surveys, interviews, and classroom observations. Students' achievement in writing was measured by scores on writing samples drawn from the district's writing assessment program and independently scored using an analytic rubric and a holistic rubric at the national scoring conference.

The analysis of teacher practices as reflected in the surveys and interviews suggests little difference between program and comparison teachers in this district in terms of writing pedagogy, knowledge, and the implementation of a writing-focused pedagogy in their classrooms. Teachers in both groups emphasized vocabulary development and made literacy resources readily available in the classroom. However, qualitative differences were consistently observed between the classrooms of program and comparison teachers in three areas: 1) program teachers were observed engaging students in a wider range of writing tasks, 2) program teachers designed writing instruction that extended over multiple class periods (sometimes lasting weeks or months), and 3) program teachers explicitly modeled reading/writing connections for their students. In contrast, writing instruction observed in comparison teachers' classrooms largely centered on completing tasks in shorter blocks of time, mirroring the type of timed writing found on state tests.

Program students' writing scores increased more than those of comparison students on both the holistic assessment and all six analytic measures. With the exception of "word choice," the differences between program and comparison students' scores were statistically significant on all measures. Furthermore, program students' scores on the Gates-MacGinitie reading test also increased more than comparison students' scores, and these differences were also statistically significant.

This study raises some important issues for sites wanting to implement a professional development program based on writing pedagogy. The statistically significant gains made by program students point to the three observed areas of difference between program and comparison teachers' classrooms—range of writing tasks, duration of writing tasks, and an explicit reading/writing connection—as aspects of writing pedagogy worthy of exploration through professional development. The reading gains made by program students suggests that

professional development based on writing pedagogy will not necessarily detract from—and may even support—instruction in reading. Finally, the similarities in knowledge of writing pedagogy between the program and comparison groups (due perhaps to an already strong focus on writing in this particular district) point to the need for future studies to find ways to discern often subtle differences in classroom practice in order to account for systematic differences in student achievement.

2. Mississippi Writing/Thinking Institute, Mississippi State University:

The Effect of Mississippi Writing/Thinking Institute Professional Development on the Writing Achievement of Ninth-Graders

This study, by the Mississippi Writing/Thinking Institute (MWTI) investigated the impact of MWTI professional development on teacher practices, to determine to what degree these professional development programs for teachers influenced student growth in writing. MWTI and its seven university-based sites provide intensive, multiyear support for schools and districts in the form of professional development workshops, classroom demonstrations, study groups, and individual teacher coaching and mentoring. This study compared teacher practices and ninth-grade-student outcomes in two MWTI program schools with those of teachers and students in two closely matched comparison schools in the 2004–2005 school year.

Using data collected through interviews and classroom audits, the study compared teacher practices of five of the six program school teachers with those of the five comparison school teachers. Implementation levels of individual strategies as well as of an aggregate of strategies were described and compared between the two groups. Holistic assessments of individual teachers' practices were also reported and compared. To measure student outcomes, the study employed a pre/post assessment design comparing writing assessment scores of 298 ninth grade students in two program schools in which MWTI had conducted professional development with scores of 157 ninth grade students in schools in which the MWTI had not provided professional support. Writing scores were compared between program and comparison students on pre- and post-instruction assessments, and within the groups for growth over time. Student writing was scored at the national scoring conference using the modified 6+1 Trait Writing Model.

The study found that program teachers implemented the teaching strategies presented during the MWTI professional development. Although statistical differences in their practices as compared to that of the comparison teachers was evidenced by only two strategies—student choice and peer response—the program teachers' interviews reflected a tendency to follow most MWTI teaching practices at a somewhat higher level of use than did the comparison teachers.

Program students outperformed comparison students in all areas. The study revealed statistically significant differences between program students' scores from pre- to posttests on holistic and all analytic measures. Comparison students, in contrast, showed little or no growth between pre- and posttests. Further, the study found that students in the two program schools that had received multiple years of MWTI professional development scored higher even on the pretests than comparison students, and then demonstrated increased growth between pre- and posttests, while their counterparts showed little or no growth. The strength of the statistically significant gains

between pre- and posttests for the program group indicates that the professional development for teachers influenced student growth in writing.

An important implication of this study is that teachers do not need to implement the full range of MWTI philosophy and strategies in order to make a significant difference in student writing; implementing individual strategies, such as peer response or student choice, may have a positive effect on student writing. The higher pretest scores of students in the schools where MWTI strategies had been implemented in middle school, coupled with students' greater gains in writing growth, further suggest that MWTI approaches may make the greatest impact on students' writing when they are learned and practiced cumulatively, over time. This suggests that more MWTI work with teachers above and below the targeted grade level could result in greater long-term gains in student writing ability.

3. New York City Writing Project, Lehman College, City University of New York: *Teacher and Student Outcomes of a Professional Development Model for Improving the Teaching of Writing*

This study investigated the impact of a partnership between the New York City Writing Project (NYCWP) and six high schools in a large urban district where challenges include poverty, low student achievement, inexperienced teachers, and increasing demands for high-stakes testing. NYCWP had worked from one to five years in each of these schools, offering support through teacher-consultants and graduate seminars, both on-site. This report presents findings from Phase One of a two-year study, conducted in 2004–2005.

The research examined how the NYCWP model of professional development supports teachers' growth and, consequently, how it affects student writing outcomes, particularly for students who face challenges when writing in English. The study compared data from two sets of teachers and their students in ninth through twelfth grade. Eight program group teachers participated in NYCWP programs at their school site; three comparison group teachers—from a demographically similar school—did not have access to NYCWP professional development. The students of teachers from both groups were more than 90% African American or Hispanic; ELL students made up 38% of the program group and 46% of the comparison group. Analysis of teachers' growth relied primarily on interviews and surveys about instructional practices and attitudes. Program and comparison group students' growth in writing was measured by pre and post assessments of student writing samples, written to prompts drawn from an archive of established writing prompts. The samples were independently assessed at the national scoring conference.

Program teachers reported adopting views and practices drawn from NYCWP professional development. They called upon the NYCWP teacher-consultants to address specific needs, and tried new strategies in their classrooms. They reported treating writing as a process, employing prewriting techniques, and engaging students in responding to each other's writing. Program teachers in content areas other than English reported incorporating writing into their instructional practice. Comparison teachers, on the other hand, felt that their professional development was less useful in their classroom practice, and employed writing strategies less consistently. Six case

studies (of four program and two comparison teachers) are presented in NYCWP's report to illustrate the effect on teacher growth of the site's work.

Program students' writing scores increased more than comparison students' scores on the holistic and all six analytic measures. In six of the seven measures ("sentence fluency" being the only exception), all differences between program and comparison students' growth were statistically significant. Subgroup analyses for ELL students demonstrated a similar pattern: program students' scores increased more than comparison students' on all measures. With the exception of "sentence fluency" and "word choice," differences between program and comparison ELL students were statistically significant. In the comparison group, with the exception of "word choice" both ELL and non-ELL students' scores either stayed the same or decreased between pre and post assessments.

The researchers caution that the small number of comparison teachers makes it difficult to draw inferences about the extent and manner of using writing in non-NYCWP classrooms. Nonetheless, the significant growth in program students' writing points to some important implications of this study. The increase in writing scores of students in history and science classes validates the programmatic effects in those contexts, and further suggests that using writing across the curriculum itself might have a positive effect on students' writing achievement. The increase in ELL program students' scores, even when students were taught by less-experienced teachers, suggests that focusing on writing as a process, and introducing writing strategies that take into account students' learning needs, can benefit all students, including those for whom writing in English is a challenge.

4. South Coast Writing Project, University of California, Santa Barbara:

Evaluating IIMPac: Teacher and Student Outcomes Through a Professional Development Program in the Teaching of Writing

This study conducted by the South Coast Writing Project (SCWriP) at the University of California, Santa Barbara examined the effects of the site's professional development program IIMPac, which focuses on the teaching of writing. IIMPac is an acronym for the five elements of the program's professional development: inquiry, inservice workshops, models, practice, and coaching. IIMPac has operated for five years in partnership with low-performing schools that serve low-income populations with a substantial proportion of English language learners.

The study focused on eight language arts teachers of fourth through eighth grade and the students in their classes, from four elementary and middle schools that were participating in IIMPac. Seven matched teachers and their students, drawn from schools with similar achievement and demographics but not participating in IIMPac, provided comparison data. Multiple measures were used to assess the effects of participation in IIMPac on teachers' classroom practices; these included surveys, interviews, classroom observations, and collections of teacher and student work. The impact of teachers' participation in the program on their students' writing performance was measured by timed writing prompts drawn from an established archive of writing assessments, and independently scored at a national conference. The impact of the program on students' attitudes toward writing was measured by a writing apprehension test. Both student assessments were administered in a pre/post fashion, in Fall 2004 and Spring 2005.

In terms of teacher practices, the study found several clear differences between program and comparison teachers. Although program and comparison teachers both taught their students strategies for preparing to write, program teachers more frequently showed their classes student-written models of writing, gave students more opportunity to develop their own writing topics, and employed a broader range of prewriting strategies. Program teachers also encouraged students to develop their ideas by revising their writing more substantially, and found ways to engage students more deeply in writing as a process of communication, including discussing their writing with other students and collecting their work in portfolios. Program teachers also facilitated student engagement in writing and ownership of written work by providing opportunities for students to share their writing, publish beyond the classroom, and amass a body of work over time.

Program students' writing performance improved more than comparison students' when measured holistically, as well as on all six of the analytic measures; however, the differences between the two groups were not large enough to be statistically significant, perhaps largely because of the general sophistication of the comparison teachers as well as the relatively high achievement of the comparison students at the outset of the study. Program students did increase their use of prewriting strategies, such as prewriting and annotating a text, to a statistically significant degree when compared with comparison students.

This study raises several important issues. First, the researchers note the challenge of matching program and comparison groups in such studies. Because comparison students began the study at a higher level of performance than program students, and because the program group included a higher percentage of ELL and low-income students, it was difficult to draw clear conclusions based upon comparative student performance. Furthermore, despite similarities in the approaches of program and comparison teachers, the study also found subtle but important differences between their practices, suggesting the importance for researchers of developing more measures of teacher practice that can capture such differences. Finally, as program teachers were relatively sophisticated in their approaches to writing instruction at the outset of the study, the study shows that the types of professional development offered by IIMPaC can help even experienced, proficient teachers improve their teaching of writing.

5. Southern Nevada Writing Project, University of Nevada, Las Vegas:

Las Vegas Through the Lens of the Family Writing Project: The Southern Nevada Writing Project's Impact on Student Writing and Teacher Practices

This study investigated the effects of participation in the Family Writing Project (FWP), a family literacy program offered by the Southern Nevada Writing Project (SNWP). The FWP serves students, parents, and teachers who live in a school district that is the fastest growing in the nation and therefore has an increasingly diverse and mobile population. FWP teacher-facilitators—who receive professional development through the SNWP—engage students and their family members in writing and art activities, discussions, publishing, and community projects. The FWP thus aims both to influence student writing (and attitudes toward writing), and to provide professional development for its teacher-facilitators.

This study sought to determine the effects of participation in the Family Writing Project on 1) student achievement in and attitudes about writing and 2) teacher practices, including approaches to the teaching of writing and relationships with students and families. Twenty-one matched pairs of middle school students (sixth through eighth grade) made up the program and comparison groups. Program students had participated in the FWP and were drawn from the classrooms of four FWP teacher-facilitators—who also made up the sample of teachers in the study. Comparison students were drawn from a school with no FWP involvement. Student achievement was measured by pre/post assessments of writing samples in response to a timed prompt, drawn from a national archive of established assessments and scored independently of the local site. Student attitudes toward writing were assessed by surveys and interviews. Analysis of teachers' classroom practices and their relationships with families drew on interviews with teachers and parents, as well as responses from student surveys.

Program students demonstrated greater growth in writing than did comparison students, as measured by their scores on each of the six analytic traits as well as the holistic score on a repeated measures ANOVA. In one area—"word choice"—the difference between the two groups was statistically significant. Furthermore, when attitudes toward writing were compared, FWP students' attitudes held steady or became more positive, while the attitudes of comparison students toward writing became less positive.

Program teacher-facilitators reported changes in their classroom practice that included improving their relationships with students and their families, deepening their understanding of how writers write, and developing their leadership and teaching skills. One of the most important implications of the data derived from teacher interviews is the importance of teachers' valuing and establishing strong personal relationships with their students and students' families. By doing so, teachers in this study better understood factors affecting their students' performance, found ways for students to draw on their own cultural and family experiences in their writing, better recognized students' strengths and assets, and supported a climate of mutual respect in their classrooms.

V. LESSONS LEARNED

A. LSRI Design Changes from Year One to Year Two

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 as it supported writing project sites in framing the most salient questions, articulating the expected effect on student learning, and identifying the most appropriate districts, schools, or classrooms to study.

Following the first year of LSRI, NWP improved the overall design of these research efforts in four ways:

- **Timing of the research award notification.** The application and notification process for NWP's second panel of LSRI sites took place substantially earlier, prior to the beginning of June. Researchers recruited participating teachers and students earlier and, therefore, were able to conduct true baseline assessments at the beginning of engagement in the program (i.e., school year or semester). Given the greater time for preparatory activity, sites were able to secure more carefully identified and appropriate comparison groups.
- **Technical assistance for research teams.** Senior NWP research staff provided increased technical assistance to the research teams at the sites. All five studies employed quasi-experimental comparative designs, and technical assistance focused on the design requirements of such research, as well as analytic issues as they pertain to both quantitative and qualitative data.
- **Field-tested measures.** A nationally developed archive of field-tested measures of student writing performance and classroom practice helped to ensure sound writing assessment across the group of studies.
- **National scoring of student writing, conducted independently of sites.** This scoring procedure ensured technical rigor and credibility of the writing assessment results. Significant improvements in the sensitivity and reliability of the scoring system permitted more powerful detection of program effects.

B. Research Design Challenges

Evaluating the effects of complex, large-scale professional development endeavors such as those of NWP will continue to require attention to two major design challenges:

- **Forming appropriate comparison groups.** Identifying, recruiting, and maintaining appropriate comparison groups continued to be quite challenging despite significant design improvements in the second panel of LSRI studies. Across diverse geographic locations and states and differing local policy environments, the LSRI experience suggests that potential comparison groups may vary considerably, with some relatively pristine and others quite sophisticated with regard to the teaching of writing. When

matching is carried out at the local level, the comparison groups necessarily vary with regard to that sophistication. The LSRI experience convinces us that in many cases the level of latent knowledge and skill among comparison-group teachers can be quite high. As a result, assessing the comparative effects of NWP programs is likely to remain a matter of comparing those programs to “everything else” that is typically occurring in the environment, where “everything else” constitutes a heterogeneous set of programs in which the level of exposure to professional development and related ideas and practices varies considerably within comparison schools and districts.

- **Detecting and describing differences in classroom practices with regard to the teaching of writing.** Closely related to the challenge of establishing appropriate comparison groups (and measuring their initial skill and sophistication) is the challenge of detecting and describing differences in classroom practice. For the most part, teachers’ self-reports on surveys and classroom-practice inventories did not reveal systematic differences in their approaches to teaching writing. Language related to the writing process appears to be ubiquitous in most teachers’ talk about writing. Self-report data would suggest widespread involvement in practices such as engaging students in prewriting activities, having them produce first drafts, and having them edit their work. However, careful analyses of interviews, of classroom observations, and of collections of writing assignments and related student work point to more nuanced differences in practice. Such nuanced measures invariably require a considerable investment, and when they involve qualitative data, analyses of their results are often difficult to aggregate across classrooms and do not lend themselves to the kind of brief summary statements required by policy documents, or to the quantification needed for incorporation into statistical analyses.

C. Opportunity to Add Expertise and Knowledge to the Field

- **LSRI provided a significant opportunity to build writing project leaders’ capacity to conduct research about the impact of their work on student outcomes.** Building such capacity is important because it allows sites to develop rigorous, research-based evidence examining in support of the services that they provide. Building research capacity also provides a forum for sites to examine the merit of the curricular and instructional ideas they advocate, as well as the efficacy of their means of developing them. Beyond that, sites are exploring ways in which the results of their research can be used strategically to advocate best practice in instruction and to collaborate with district and school administrators to develop effective inservice partnership arrangements. An overall effect of the research initiative is the expansion of the pool of program developers and researchers with expertise in evaluating the effect of professional development programs on the teaching of writing.
- **The large-scale scoring of student writing provides an opportunity to add to our professional knowledge base by exploring related basic research questions.** The LSRI’s national scoring conference, at which student writing samples from all sites were scored, afforded an opportunity to examine a number of research questions related to the assessment of writing. Specifically, two studies were conducted:

The effects of scoring order on holistic and analytic writing assessment scores.

Conventional wisdom suggests that to preserve the independence of holistic judgments they should precede analytic scoring. However, little is known about the effects of scoring order or if true holistic scoring is even possible when a scorer also provides analytic scores. Student writing was scored under carefully controlled conditions (holistic then analytic; analytic then holistic; pure analytic; and pure holistic) to explore the effects of scoring order on the nature of scores that each condition provides.

A comparison of full-scale and forced-choice scoring methods. Typically, writing is scored by comparison to anchor papers that exemplify values on a full scale of four to six points. An alternative proposed for comparative questions such as evaluations requires scorers to read two masked writing samples, choosing one or the other as “better.” This forced-choice approach is proposed as more economical, more reliable, and more sensitive to modest differences than full-scale scoring. Papers were scored using both approaches to allow comparison of the two.

Each of these studies will be written up for professional publication, thus contributing to the professional knowledge base about the assessment of writing. In addition, the knowledge gained from these studies will also be used to inform design and assessment decisions as the NWP defines a long-term national evaluation of the impact of its programming.

VI. CONCLUSION

The second year of LSRI has contributed in several ways to NWP’s national research and evaluation agenda. First, methodological improvements made the studies better able to detect differences in performance between program and comparison groups when they existed. Because of the increased methodological rigor, this research adds a series of sharply focused studies to the NWP’s national evaluation portfolio. Those studies uniformly suggest the positive impact of the NWP programming that they investigated. Second, the LSRI research teams report enhanced research skill and knowledge because of their engagement in this work. The NWP is building a national pool of researchers and program leaders with expertise in evaluating professional development programs in the teaching of writing. Third, the LSRI’s national scoring conference afforded an opportunity to explore technical questions related to the assessment of writing. In this way, the LSRI goes beyond the typical contribution of evaluation studies and promises to add to the national knowledge base about the assessment of writing.

In summary, this year’s LSRI subsumes carefully conducted studies from five NWP sites in diverse geographic locations from around the nation. These quasi-experimental studies indicate positive effects for the students of teachers who participate in writing project programs. Every single comparison across the five studies gives evidence of the positive effects of the NWP programming. Four of the five studies demonstrated statistically significant differences that

avored the program group on at least one measure of student writing performance. Student results were strong and favorable in those aspects of writing that the NWP is best known for, such as the development of ideas and organization. At the same time, students in writing project classrooms outgained their peers in the area of conventions as well, suggesting that even these basic skills benefit from the NWP approach to the teaching of writing.

VII. REFERENCES

- Bellamy, P.C., ed. 2005. *Seeing with New Eyes*. Portland, OR: Northwest Regional Educational Laboratory.
- Culham, R. 2003. *6+1 Traits of Writing*. New York: 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.
- National Research Council. 2002. *Scientific Research in Education*. Washington, D.C.: National Academy Press.
- National Writing Project. 2005. *Local Site Research Initiative Report, Cohort I (2003–2004)*. Berkeley, CA: Author.
- Northwest Regional Educational Laboratory. 1998. *An Introduction to 6+1 Trait Writing for Assessment and Instruction*. Portland, OR: Author.

**Appendix A:
Inter-rater Reliabilities by Trait and by Level**

Level	Total N of Papers	Number of papers double scored	Double Score Rate	Number of contrasts	Total # adj.	Total % agree	Holistic # adj.	Holistic % agree	Ideas # adj.	Ideas % agree
Elementary	1188	136	11%	952	101	89%	7	95%	14	90%
Middle School	2379	1231	52%	8617	614	93%	54	96%	95	92%
High School	3938	381	10%	2667	289	89%	23	94%	39	90%
Total - all levels	7505	1748	23%	12236	1004	92%	84	95%	148	92%

Level	Organization # adj.	Organization % agree	Voice # adj.	Voice % agree	Sentence Fluency # adj.	Sentence Fluency % agree	Word Choice # adj.	Word Choice % agree	Conventions # adj.	Conventions % agree
Elementary	13	90%	16	88%	18	87%	16	88%	17	88%
Middle School	94	92%	107	91%	99	92%	68	94%	97	92%
High School	42	89%	54	86%	48	87%	31	92%	52	86%
Total - all levels	149	91%	177	90%	165	91%	115	93%	166	91%