

**Evaluation Report:
Effectiveness of the North Carolina State Improvement Project in
Improving the Reading Performance of Students with Disabilities**

Executive Summary

This report provides information and data to evaluate the effectiveness of the North Carolina State Improvement Project (NC SIP) in improving the basic reading skills performance of students with disabilities. Data were collected on the implementation of reading instruction and the reading performance of students during the 2001-2002 school year. The centers' instructional strategies and procedures were designed to reflect the principles of reading instruction derived from an extensive body of scientifically sound research studies. These principles embrace the use of direct, explicit, and systematic instructional techniques and strategies.

Data were collected on a non-random sample of 200 students receiving instruction in the eight reading centers. However, missing data reduced the number of students included in the various data analyses. Non-participation of some students in the test administrations was the primary cause of missing data.

Using the test results from 2000-2001 and 2001-2002 end-of grade reading tests, a comparison of the progress in reading performance of students receiving the NC SIP reading instruction with the progress in reading performance of all students with disabilities in North Carolina was conducted. These data indicate that students with disabilities receiving reading instruction through the North Carolina State Improvement Project gained at a rate four times greater than students with disabilities statewide. These results support a conclusion that the reading instruction as implemented in the NC SIP Best Practice Centers is extremely effective for students with disabilities. Results of analyses of a pre- and post-instruction administration of the reading subtests from the Woodcock Johnson III Test of Achievement support this conclusion.

During the next two years efforts will be made to increase the number of students receiving the services available through North Carolina State Improvement Project. Evaluation of the reading progress of students enrolled in the North Carolina State Improvement Project will be expanded to include more students and to improve the reliability of the data collection procedures.

I. Introduction

This document reports on the effectiveness of the North Carolina State Improvement Project (NC SIP) in improving the basic reading skills performance of students with disabilities. To address this goal eight Best Practice Centers have been established in public school systems across North Carolina. The Centers are designed to implement and demonstrate the use of research-proven instructional strategies for teaching reading to students with severe and persistent reading problems. The Centers were initiated in the fall of 2000 and the first year was devoted to staff-development and the implementation of their instructional programs.

During their second year, the centers continued to receive training and technical assistance as they expanded their instructional programs and participated in the NC SIP evaluation program.

The NC SIP focus on reading instruction has been stimulated by three primary considerations. First, the reading and writing performance of students with disabilities enrolled in the regular curriculum is dismal. Students with disabilities participating in the state's accountability testing program score approximately at one-half the reading level of the total population of students taking the tests. The consequences of this low level of performance are striking. Students not performing at or above grade level may be: (a) held back at grade level, (b) denied a diploma, or (c) shifted from a diploma curriculum track to a non-diploma curriculum track. These alternatives are associated with school drop out rates for students with disabilities that are substantially higher than the rate for students without disabilities.

Second, there is clear and specific research evidence that the vast majority of students with disabilities can learn to read on grade level, if appropriate research-validated instruction and learning techniques are effectively employed.

Third, the new Federal No Child Left Behind (NCLB) law requires that the basic skills progress of students with disabilities be disaggregated and reported. Students with disabilities enrolled in the standard curriculum must meet the adequate yearly progress guidelines established by the U.S. Department of Education for all groups of students. Schools that do not meet the State established standards for adequate yearly progress across all groups of students, including students with disabilities, will face penalties and/or sanctions. Currently the yearly progress of students with disabilities in North Carolina on performance in basic academic skills is lower than any other designated group to be evaluated as required in NCLB.

II. The NC SIP Reading Instruction Program

The NC SIP staff development program to improve reading instruction has been planned to reflect the basic principles gleaned from an extensive body of instructional research that includes students with reading difficulties. The program reflects the findings of two recent National reports: *Teaching Children to Read*, a report of the National Reading Panel, and *Preventing Reading Problems of Young Children*, a report sponsored by the National Reading Council of the National Academy of Sciences. In addition, the program reflects the findings of several studies of the type of instruction needed to effectively remediate students who have failed to learn to read. The principles of reading instruction gleaned from the research and reflected in the NC SIP training include the use of direct, explicit, and systematic instructional techniques and strategies.

The content in the staff development program was developed by Rebecca Felton and David Lillie, in partnership with the Guilford County Schools. Video media included in the training CD used in the project is used with permission of the Guilford County Schools. The training program consists of twelve units as listed below.

1. Introduction
2. Learning to Read, A National Problem
3. The Major Principles of Reading Instruction
4. The Structure of Language - What Teachers Need to Know
5. Assessment of Basic Reading Skills
6. Teaching Phonemic Awareness
7. Teaching Letter-Sound Associations
8. Teaching Word Identification: Decoding and Sight Words
9. Teaching Spelling
10. Developing Automaticity and Fluency
11. Teaching Reading Comprehension
12. Selecting and Implementing an Effective Reading Program

In addition to staff development that clearly reflects research-proven instructional strategies, NC SIP has monitored the fidelity of the implementation of instruction in each center and has provided continuous technical assistance and support. A detailed description of the staff training, technical assistance, and evaluation assistance provided by the NC SIP can be found at www.ncsip.org.

III. Evaluation Procedures

Data were collected during the 2001-2002 school year to provide information to document the (a) characteristics of students participating in the evaluation, (b) characteristics of instructional settings across centers, (c) level of fidelity of the centers' instructional model, and (d) improvement of students' reading skills using the 2000-2001 and the 2001-2002 scaled scores from the North Carolina End-of-Grade reading tests. Also pre-and post-test data were collected using the reading subtests of the Woodcock-Johnson III Tests of Achievement. The SAS JMP statistics software program was used to conduct the statistical analyses.

Each of the NC SIP Best Practice Centers collected evaluation data on those students who were receiving instruction from teaching staff identified as "model instruction" teachers. Each of the centers is in the process of expanding the use of their model reading instruction program within their own school district. Only students receiving model reading instruction from teachers who completed the NC SIP Research to Practice training as well as additional training in the implementation of the center's specific model instructional program were included in the evaluation population. In addition to the students involved in the evaluation effort, each center impacts a larger number of students than reflected in the evaluation population.

A. Characteristics of the Student Sample

Characteristics of the students participating in the NC SIP evaluation program are presented in Table 1 below. Initially 200 students enrolled across the reading Best Practice Centers were identified for the evaluation sample. However, missing data reduced the primary

evaluation sample (when analyses included EOG scores for 2001-2002) to 133 students. Non-participation of students in the sample in one or both of the EOG reading test administrations was the primary cause of missing data.

Table 1: Characteristics of the Student Sample

Identification	Gender	Ethnicity	Grade
BED - 1	F - 33	White - 80	G3 - 16
EMD - 4	M - 100	Black - 44	G4 - 24
SLD - 107		Hispanic - 8	G5 - 17
OHI - 12		Other - 1	G6 - 32
Other - 9			G7 - 20
			G8 - 16
			G9 - 2
			G10 - 2
			G11 - 3
			G12 - 1
Total 133	133	133	133

As can be seen in Table 1, 75% of the students included in the evaluation study were identified as having a specific learning disability. Twelve students were identified as Other Health Impaired, which includes students with Attention Deficit Disorders. Other disabilities included students with physical impairments, traumatic brain injury, and trainable mental impairment. All students included in the program are receiving the North Carolina Standard Curriculum, with approximately 75% Male, 60% White, and 33% Black students.

B. Data Collection Procedures

Each of the eight NC SIP Best Practice Centers was provided with instructions for collection of evaluation information. Data and information were collected across the three categories of (a) student characteristics, (b) dimensions of the instructional situation, and (c) student reading performance. Specific data types are listed below.

1. Student Characteristics: Grade, Gender, Ethnic Background, and Type of Disability.
2. Dimensions of the Instructional Situation: Instructional Fidelity Observations, Instructional Setting, Group Size, and Number of Instructional Sessions.
3. Student Reading Performance: Student's Scaled Scores on the ABC End-of-Grade Reading Tests for the 2000-2001 and 2001-2002 years, and Pre- and Post-Test Woodcock Johnson Psycho-Educational Battery Reading Achievement Subtest and Cluster Scores for each student.

C. Evaluation Procedures Limitations

It should be noted that the procedures used in the NC SIP evaluation study do not reflect the rigor and control that characterize scientific research. Data were collected on students already

placed in instructional groups. Random selection of students with disabilities for placement in the NC SIP project classroom was not employed. Students included in the evaluation study were selected to participate in the reading center programs because of their lack of progress in reading achievement and their need for intense and explicit reading instruction. In addition, the reliability of some of the data collection instruments and procedures used was not established. Finally, analyses including all students in the initial evaluation plan could not be conducted due to incomplete data for some students.

IV. Evaluation Results and Discussion

The NC SIP evaluation findings are presented across several topics that include (a) progress in reading as measured by student EOG reading performance, (b) progress in reading as measured by student performance on the repeated administration of the reading subtests of the Woodcock Johnson III Tests of Achievement, and (c) instructional setting factors and their relationship to reading progress of students.

A. Reading End-of-Grade Tests Results

Below, Table 2 presents the comparison of the reading performance progress of students receiving NC SIP reading instruction with the reading performance progress of all students with disabilities included in the North Carolina ABC EOG reading testing. This comparison uses the ABC Accountability test results from 2000-2001 and 2001-2002.

Table 2: Comparison of Reading EOG Mean Scores and Gain Scores of North Carolina Students with Disabilities and NC SIP Students From 2000-2001 to 2001-2002

		2001		2002	
Population	N	Mean SS	N	Mean SS	Gain
North Carolina	67804	147.4	66490	148.7	1.3
NC SIP	133	141.6	133	146.9	5.37

Note: North Carolina data represent student populations that changed from 2001 to 2002. The NC SIP data represent the same students from 2001 to 2002.

These data indicate that from the 2001 spring testing to the 2002 spring testing students with disabilities receiving reading instruction through the North Carolina State Improvement Project gained at a rate four times greater than students with disabilities statewide. The gain scores across the eight NC SIP centers ranged from 1.17 to 11 with a mean gain score of 5.37.

It is interesting to note that in the spring of 2001, NC SIP students scored, on the average, 5.8 points lower than the statewide population of students with disabilities participating in the EOG reading testing. It is assumed that this difference is related to a selection factor bias resulting when students with the most severe and persistent reading problems were selected first for participation in the project. After receiving instruction for one academic year, the sample of students in the NC SIP project scored, on the average, only 1.8 scale score points lower than the

statewide population of students with disabilities. These results indicate that the NC SIP students gained at a rate of progress four times greater than the rate of progress of the statewide population of students with disabilities. The extent to which this rate of progress continues as students are maintained in the project will be explored as data are collected in 2003 and 2004.

B. Woodcock Johnson III Achievement Tests Results

In addition to the comparison of reading EOG scores, reading subtests from the Woodcock Johnson III Test of Achievement (WJ III ACH) were administered in a pre- and post-test fashion during the 2001-2002 school year. The WJ III ACH standard scores are based on a mean of 100 and standard deviation of 15. A student achieving at an average level for his or her grade level will demonstrate a standard score in a range of 90 through 110. Gains in standard scores indicate progress above and beyond that expected from year to year. As can be seen in Table 3, gains in standard scores were made across all the reading subtests administered.

Table 3: Pre- and Post-Test Standard Scores, Means and Mean Gains of Students Enrolled in the NC SIP Reading Centers Across Woodcock Johnson III Test of Achievement Reading Skills Subtest Areas 2001-2002.

Reading Skills	N*	Pre-test mean	Post-test mean	Gain
Letter-word	185	77.43	80	2.57
Word Attack	185	84.02	88.81	4.79
Spelling Sounds	148	85.86	90.65	4.79
Phoneme Know.	148	84.17	89.30	5.13
Basic Rdg. Skills	141	80.67	84.71	4.04

* Not all sub-tests were administered to the total sample.

The largest gains from pre-testing to post-testing were made on the Phoneme Knowledge subtest, which, on the average, improved to within one standard score point of an average level of performance. Spelling of Sounds subtest performance improved to an average level of performance. These average gains in standard scores are impressive since they indicate that students receiving instruction in the NC SIP centers have, on the average, improved their reading abilities at a faster rate than expected of the average student. It should also be noted that for most of the students included in the evaluation study the period of instruction between the pre-test and post-test was substantially less than a full academic year of instruction. Some students were not administered the pre-test until January, after several months of participation in the SIP instructional program.

C. Relationships Between Instructional Setting Variables and Student Progress

To investigate the relationship between instructional setting differences across centers and the progress in reading performance as measured by the reading EOG scores from 2000-2001 and 2001-2002, data were collected on several instructional setting variables across the eight NC SIP centers. The instructional variables included:

- (a) Setting: The type of classroom instructional setting in which the student is receiving the reading instruction including Self-Contained Classroom, Resource Room, and/or a General Education Classroom.
- (b) Fidelity: A composite rating of the level of fidelity to the instructional model being implemented as measured through multiple classroom observations of the implementation of reading lessons in the classroom.
- (c) Group Size: The number of students in the reading instruction group.
- (d) Instructional Sessions: The number of instructional sessions each student received.

Using a visual review of the “setting” information, there appears to be no observable relationship between the type of setting (Self-Contained Classroom, Resource Room, and General Education Classroom) and student progress in reading achievement. To investigate the relationships between the instructional setting variables of fidelity, group size, and number of instructional sessions; data on each of these variables were analyzed using the Spearman Rho measure of correlation coefficient. These results are shown in Table 4.

Table 4: Rank Order Correlations Among Instruction Variables and Reading Gains

	Variable	By Variable	Spearman Rho
1	Group Size	Fidelity	-0.4857
2	No. of Inst. Sessions	Fidelity	-0.0857
3	No. of Inst. Sessions	Group Size	0.7714
4	Reading Gain	Fidelity	-0.1429
5	Reading Gain	Group Size	-0.3952
6	Reading Gain	No. of Inst. Sessions	-0.4286
7	Reading Gain	Developmental Profile	0.3952
8	Developmental Profile	Fidelity	0.6000
9	Developmental Profile	Group Size	-0.3571
10	Developmental Profile	No. of Instr. Sessions	0.5429

It should be noted that these analyses are exploratory in nature. Before these data can be used with confidence, the reliability of the data collection procedures must be established. This has been done for the fidelity variable, but not for the other two instruction variables. Although reporting group size and number of instructional sessions appears to be an easy task of counting and record keeping, no procedures were used to verify the accuracy of the numbers reported.

Procedures were conducted to determine the reliability of the data collection procedures used to determine the extent of instructional fidelity. The average model fidelity score for a center was determined using a structured classroom observational rating system developed to measure the extent to which a center teacher delivered the reading lesson as designed and evaluated by the developers of the instructional model being used by the center (*SRA Corrective/Reading Mastery, Language!, Wilson Reading System, or LiPS*). The rating scales used to measure instructional fidelity can be found on the NC SIP website (www.ncsip.org) in the Accountability and Evaluation section. Each teacher was observed teaching an instructional lesson by a trained observer. Most of the teachers delivering the reading instruction were

observed once a month (for four months) by the two NC SIP reading specialists and a trained observer from the Best Practice Center. For the six centers reporting fidelity scores, independent observer agreement ranged from 75% to 96%, with an average agreement of 86%.

Within the context of the limitations indicated above, the relationships among the instructional variables and the reading gains are presented below.

1. Instructional Group Size and Fidelity of the Instructional Implementation

There is a modest negative relationship (-.49) between the size of the instructional group and the fidelity of the instructional model implementation. The average size of an instructional group for a center ranged from a low of 1.17 to a high of 6.82 students in a group. In general, centers with smaller instructional groups tend to have higher instructional fidelity. There are several possible explanations for this relationship. One is that it may be easier to deliver instruction as prescribed by the model developers with smaller groups of students. The NC SIP project will continue to investigate this possibility.

2. Number of Instructional Sessions and Fidelity of the Instructional Implementation

There is no meaningful relationship between number of instructional sessions and fidelity. Regardless of the average number of instructional sessions provided by a center the instructional fidelity could be low or high.

3. Number of Instructional Sessions and Instructional Group Size

There is a substantial relationship (.77) between the number of instructional sessions reported by the centers and the size of the instructional group. This relationship indicates that, in general, as the number of instructional lessons in which a student participated increased, the size of the instructional group increased. The average size of an instructional group for a center ranged from a low of 1.17 to a high of 6.82 and the average number of instructional lessons received by a student in a center ranged from a low of 69 to a high of 144. One possible explanation for this relationship is that when the reading instruction was integrated into the existing instructional schedule of a school, intact groups of students were used, as opposed to pullouts or individual tutoring which are more difficult to schedule on a routine basis.

4. Fidelity of the Instructional Implementation and Gains in Reading Performance

There is no meaningful relationship between the fidelity of the implementation of the instruction and the reading gains of students. This finding could be related to the nature and design of the fidelity measurement procedures. The instructional fidelity rating scales for each model use an ordinal criterion scale where the highest score of three indicates an “appropriate” level of instructional fidelity. A rating of 1 indicates that the instructional fidelity is “inappropriate” and a rating of 2 indicates instruction that is “somewhat appropriate.” The average fidelity ratings across the centers range from a high of 2.98 to a low of 2.49. Thus, teachers in all the centers reporting fidelity ratings were observed to be delivering instruction

well above a “somewhat appropriate” range. Therefore, it is possible that a center with an average fidelity rating of 2.49 may be implementing the instruction at a level that is equally effective as a center with a rating of 2.98.

5. Instructional Group Size and Gains in Reading Performance

There is a modest negative relationship (-.40) between instructional group size and reading performance gains of students. This relationship does not mean that all that is necessary to produce higher reading gains is to reduce the size of the instructional group. However, it is reasonable to conclude that smaller instructional group size is one of several important instructional variables related to gains in student performance.

6. Number of Instructional Sessions and Gains in Reading Performance

There is a modest negative relationship (-.43) between the number of instructional sessions and the reading gains of students. While this may appear to be counter-intuitive, several factors may have contributed to this finding. One of the most effective NC SIP centers employed an instruction model that involved teaching students in one-to-one settings for fewer instructional sessions (in comparison to models that provided instruction in larger groups but on a more frequent basis). It is assumed that instructional factors other than the number of reading instruction sessions are also related to reading gains. Instructional group size and fidelity of the instructional implementation are other important considerations.

D. Program Development Ratings and Reading Gains

A rating instrument, the NC SIP Developmental Profile was developed and used to assist the project in determining the quality of each center’s program implementation. The rating profile was designed to measure the developmental level of the center across four areas of program development: (A) Implementation of Center’s Model Instruction, (B) Research To Practice Foundation Training, (C) Model Implementation Training, and (D) Student Progress Evaluation. Ratings were conducted across a series of important program tasks using a scale of 1 (No progress on this task) through 4 (Yes, task completed or demonstrated in exemplary fashion). A copy of the NC SIP Developmental Profile and instructions for its use can be found on the NC SIP website www.ncsip.org in the evaluation and accountability section of the website. In most instances, the same NC SIP personnel were involved in the site-visit procedures used to rate the developmental level of each center.

As can be seen in Table 4 on page 7, modest relationships exist between the developmental profile ratings and gains in reading (.40), instructional group size (-.36), fidelity of instruction (.60), and number of instructional sessions (.54). To a modest degree the use of the developmental profile can be useful in predicting the average gains in reading of the students enrolled in the program. The developmental profile observational ratings are also, in general, related to the level of instructional fidelity, the size of instructional groups, and the number of instructional sessions. The NC SIP project will continue to use the developmental profile to assist the project with the planning of technical assistance services and additional staff training for individual centers and projects in the NC SIP network.

Best Practice Centers

Center	Focus of Center	Contact Information
South Central NC Cumberland County Joyce Carter, Coordinator jcarter@ccs.k12.nc.us	Reading and Writing: Explicit code-based reading instruction using the Wilson Reading System	Chestnut EC Office 2121 Skibo Rd Fayetteville, NC 28314
Western NC Transylvania County Kathy Haehnel, Coordinator khaehnel@transylvania.k12.nc.us	Reading and Writing: Explicit code-based reading instruction using the Wilson Reading System	Transylvania County Schools 400 Rosenwald Lane Brevard, NC 28712
Western NC Haywood County Sharon Burgin, Co-Coordinator sharonb@haywood.k12.nc.us Lynn Bailey, Co-Coordinator Lbailey@haywood.k12.nc.us	Reading and Writing: Explicit code-based reading instruction using the Wilson Reading System	Haywood County Schools 1230 North Main St. Waynesville, NC 28786
Eastern NC Onslow County Ann Spangler, Coordinator aspangler@onslow.k12.nc.us	Reading and Writing: Explicit code-based reading instruction using the Language! Literacy Intervention Curriculum Program	Onslow County Schools 200 Broadhurst Rd Jacksonville, NC 28540
North Central NC Wake County Connie Steigerwald, Coordinator csteigerwald@wcpss.net	Reading and Writing: Explicit code-based reading instruction using SRA Corrective Reading, SRA Reading Mastery, and Great Leaps	Wake County Schools 4401 Atlantic Ave Raleigh, NC 27604
South Central NC Montgomery County Sandra Miller, Coordinator sramiller@hotmail.com	Reading and Writing: Explicit code-based reading instruction using the Lips method and Language!	Montgomery County Schools Anchor Alternative School Lambert Rd Biscoe, NC 27209
Northeastern NC Northampton County Linda Thomas, Coordinator thomasl.co@ncs.schoollink.net	Reading and Writing: Explicit code-based reading instruction using the Language! Literacy Intervention Curriculum Program	Northampton County Schools P.O. Box 158 Jackson, NC 27845
Northwestern NC Watauga County Vickie Norris, Co-Coordinator norrisv@watauga.k12.nc.us Stamey Carter, Co-Coordinator carters@watauga.k12.nc.us	Reading and Writing: Explicit code-based reading instruction using the Wilson Reading System	Hardin Elementary School 361 Jefferson Rd Boone, NC 28607