

Research and Evaluation from South Carolina Reading First and Lessons Learned

Tammiee S. Dickenson

Diane M. Monrad

Robert L. Johnson

Mihaela Ene

Elizabeth Leighton

Tomonori Ishikawa

Zhaoxia Guo

University of South Carolina

Paper presented at the

Annual Conference of the South Carolina Educators for the Practical Use of Research,

Winner of the 2015 Distinguished Paper Award Competition

Columbia, SC February 26-27, 2015

and at the

Annual Meeting of the American Educational Research Association,

Invited Session for the Consortium of State and Regional Educational Research Associations

Chicago, IL April 16-20, 2015

Abstract

As South Carolina moves forward with the implementation of the recently passed Read to Succeed legislation, it is important to reflect on lessons learned from earlier reading initiatives in the state. This paper includes a synthesis of literature, evaluation findings, and research results based on the South Carolina Reading First (SCRF) model. Highlights are shared from various studies investigating the reading achievement of students who participated in the SCRF Initiative from 2004-2010 which can be used to inform school programming under Read to Succeed, the state's current focus to improve reading instruction state-wide. Lessons learned include the importance of collaborative evaluation, commitment and involvement of all stakeholders, the critical role of assessment, addressing summer loss, monitoring implementation fidelity, and establishing program outcomes and expectations.

Lessons Learned from Research and Evaluation on Reading Initiatives in South Carolina

As South Carolina (SC) moves forward with the implementation of the recently passed Read to Succeed legislation, it is important to reflect on lessons learned from earlier reading initiatives in the state. The authors of this paper have been involved in research and evaluation associated with multiple reading initiatives, including SC READS, the SC Reading Initiative, and the federally funded SC Reading First (SCRF) Initiative. This paper focuses on the SCRF Initiative and provides a synthesis of literature, evaluation findings, and research results based on the SCRF model. Key findings from this body of work are provided along with implications that can be used to inform best practices for the state's current focus on improving reading instruction state-wide.

Program Information

Read to Succeed

Read to Succeed (R2S) was passed by the SC General Assembly and signed into law by Governor Nikki Haley in June of 2014 with implementation to begin in the 2014-2015 school year. The purpose of R2S is to ensure all students gain reading and writing proficiency in order to prepare them for post-graduation success in careers and postsecondary education. Under the law, students who are unable to comprehend appropriate grade level texts will be identified as early as possible in order to receive targeted support from their classroom teacher, as well as additional support from a reading interventionist on an as-needed basis. By providing students with a comprehensive support system and intensive interventions, students are expected to attain proficiency in reading by third grade or risk being retained.

RESEARCH AND EVALUATION FROM SOUTH CAROLINA READING FIRST

The R2S law aims to promote reading and writing skills across all academic disciplines, and requires the development of reading proficiency plans by both the state and each district.

These plans are designed to highlight major components of the program including

- professional development opportunities to increase teachers' reading expertise and foster literacy leadership by coaches and administrators,
- comprehensive assessment system to be used to monitor reading achievement and growth,
- implementation of effective instructional practices to improve students' reading and writing across disciplines,
- implementation of effective intervention strategies to support struggling readers and early childhood interventions,
- support systems for families to assist children with literacy in the home, and
- preparation of early childhood teachers to provide a curriculum that is rich in language and literacy.

Districts are also required to outline partnerships with community organizations (e.g., libraries, arts organizations, etc.) to promote reading. In addition to state and district reading proficiency plans, individual schools are required to develop school reading proficiency implementation plans. These school specific plans are to be aligned with their district's plan and should serve as guidance for classroom teachers relating to the strategies being implemented to address the reading and writing needs of students in their school.

Key requirements for schools include professional development in evidence- and research-based strategies of reading instruction for teachers, the use of assessments for progress monitoring, employment of reading/literacy coaches to support teachers, the provision of

intervention for struggling readers, communication with parents regarding student progress and home supports, and the offering of summer reading camps for students at-risk for retention. In addition to ensuring all current teachers receive adequate training in the area of literacy, the R2S law also places emphasis on pre-service teacher training. The newly created Read to Succeed Office in the South Carolina Department of Education (SCDE) will work collaboratively with faculty from higher education institutions to establish “essential competencies” that all certified teachers need to know to ensure students comprehend grade-level texts. All students entering either a pre-service teacher education program or a MAT degree program in 2016-2017 must complete a specified course sequence in literacy. This includes a 12-semester credit sequence for those seeking to obtain certification in early childhood or elementary education, and a 6-semester credit sequence for those seeking certification in middle or secondary education. With these comprehensive supports in place, students are expected to achieve proficiency on the state reading assessment.

South Carolina Reading First

Reading First, part of the federal No Child Left Behind Act of 2001, was a nationwide effort to provide states and school districts with support to establish research-based reading programs for students in kindergarten through third grade. The evaluation of the SCRF Initiative was a collaborative effort between the external evaluators, the Office of Program Evaluation (OPE) and the South Carolina Educational Policy Center (SCEPC), and personnel at the SCDE. This on-going collaboration involved regular meetings and communication through which project implementation and evaluation activities were planned and results were shared. Numerous reports and presentations related to the SCRF Initiative were completed during the time period the SCRF Initiative began providing program services in the 2004-2005 school year.

RESEARCH AND EVALUATION FROM SOUTH CAROLINA READING FIRST

Evaluation results were shared in annual reports (Dickenson et al., 2005; Sesso-Dahlke et al., 2006; Bennett et al., 2007; Bennett et al., 2008; Gareau et al., 2009; Bennett et al., 2010) as well as in meetings with SCRF project staff, professional development providers, School Leadership Teams, and regional literacy specialists, who then shared the results at the school level.

Members of the evaluation team also reported results at national meetings of the American Educational Research Association and American Evaluation Association (Dickenson & Monrad, 2007; Dickenson, Smith, & Gilmore, 2007; Dickenson, 2008; Dickenson, Monrad, Johnson, & Wills, 2008; Dickenson & Smith, 2009; Dickenson, 2009; Dickenson & Young, 2010; Ishikawa et al., 2010; Dickenson, Monrad, May, & Bennett, 2011; Morgan, Dickenson & Young, 2012).

The SCRF Initiative began providing program services during the 2004-2005 school year to approximately 13,000 students in 52 schools from 24 districts in the state. The origins of the initiative trace back to the South Carolina Reading Initiative (SCRI), which was created in February, 2000 as a result of input from the first SC Reading Summit, the Governor's Institute of Reading Task Force, and a review of best practices in the teaching of reading (National Council of Teachers of English, 2008). The SCDE received its Reading First award of approximately \$90 million total funding in 2003, and spent the 2003-2004 school year in planning, initial coach training, and infrastructure development. Selection processes were used to ensure that curricula and assessments philosophically aligned with the pedagogical approach emphasized by the SCRF model.

The goal of the SCRF Initiative was *to improve reading achievement in grades K-3 so that all children are reading at the appropriate grade level*. To achieve this goal, SCRF had three objectives:

RESEARCH AND EVALUATION FROM SOUTH CAROLINA READING FIRST

- Enable and motivate teachers to understand and confidently implement scientifically-based reading research (SBRR) programs, strategies, skills, and assessments in their classrooms.
- Support the change process from the "bottom up" by supporting collaboration and conversation at various levels to ensure the sustainability of this initiative.
- Establish and expand an increasing pool of teachers and administrators who are knowledgeable about, committed to using, and successful in teaching a comprehensive reading program based upon scientific research.

Professional development model.

School-based literacy coaches with specialized training in the teaching of reading were utilized to create job-embedded professional development opportunities for teachers. The coaches were supported by regional literacy specialists and state liaisons, employed by the SCDE. The SCDE's SCRF staff engaged in program monitoring and implementation, record keeping, and delivering professional development. The teachers and administrators in SCRF schools were required to attend weekly study groups and professional development sessions focusing on strategies to teach key reading components. In addition to instructional coaching, interventionists at each school provided targeted reading instruction to the lowest performing students in either one-to-one or small group sessions. Further, SCRF School Leadership Teams (SLT), including teachers, school literacy coaches, media specialists, principals, and other school/district personnel, participated in professional development to promote collaborative decision-making and responsibility and to receive guidance in the implementation of the grant. The intent was for professional development to support the interrelationship between SCDE

RESEARCH AND EVALUATION FROM SOUTH CAROLINA READING FIRST

staff, program participants, school administrators, and other stakeholders and to contribute to program success.

The five components of reading instruction from the National Reading Panel (2000) that were emphasized by SCRF included phonemic awareness, phonics, fluency, vocabulary, and comprehension. National experts on each of the five components conducted training for state and regional personnel during the first year of SCRF as a “train the trainer” approach. The state and regional personnel provided professional development and support to the school-based literacy coaches. All literacy coaches were certified at the elementary level, held a master’s degree, and completed training relating to SBRR instruction and assessment. The literacy coaches were employed by each school to assist teachers with implementing the strategies learned in the professional development sessions. Through professional development and support from literacy coaches, the intent was for teachers to be well prepared to provide appropriate instruction that would lead to improved reading achievement for all of their students.

Progress monitoring and intervention.

At the inception of SCRF, district leadership teams were introduced to the concept of a tiered instructional delivery model that rested upon a foundation of good first teaching in every classroom. The progress of all students was monitored using three subtests from the Dominic Reading and Writing Portfolio Assessment (DeFord, 2000), which is intended for use with students in grades K-8 and assists teachers in documenting growth in achievement and instructional decision-making (Pearson Learning Group, 2004). The Dominic was administered three times during the year to all students in grades K-3 and more frequently for students identified for intervention. Students were flagged for intervention at the beginning of the year

based on low Dominie scores. Students who did not make adequate progress between administrations were identified for additional assistance through intervention.

The first and foremost component of the SCRF intervention model was the provision of quality reading instruction for all students in a 120-minute uninterrupted block. This reading block consisted of flexible grouping with instruction provided for the whole class and in small groups. The participation of teachers in study groups focused on SBRR strategies and support from literacy coaches served to strengthen instruction in all classrooms. This foundation of instruction, focused on the five reading components, was considered as Tier 1 intervention.

Types of reading intervention. Consistent with Reading First's focus on increasing the intensity of instruction for struggling readers, SCRF interventions were designed to increase instructional time, the amount students were engaged in high success reading and writing, the intensity of instruction through customization and differentiation, and teacher attention to individuals by decreasing group size to match the needs of students. SCRF interventions were coded as either Reading First Additional small group interventions if intervention was provided in small groups of 4-5 students each or Reading First Substantial if interventions were provided in groups of 1-3 students each. Reading Recovery® (Pinnell, DeFord, & Lyons, 1988) functioned as a substantial intervention, providing intensive, short-term, one-on-one tutoring for the lowest performing first graders. Students moved flexibly between levels of intervention based upon their progress as measured by the state's progress monitoring instrument, Dominie.

Training and support for SCRF interventionists. During district leadership meetings in the first year of SCRF, district teams were provided guidance in selecting and hiring reading interventionists. These teachers were required to have at least 5 five years of successful teaching experience at the elementary level, excellent problem-solving abilities, and be willing to commit

to extensive, ongoing training. SCRF offered training and support of interventionists for both one-on-one and small group interventions.

School intervention teams. Each SCRF school had a School Intervention Team (SIT) that met monthly to monitor the progress of students in interventions, manage movement of students in and out of intervention services, and match students with interventions appropriate to their needs. SIT members included the school's interventionists, literacy coach, administrator responsible for instructional decision-making and scheduling, special education teacher(s), teacher representative from first grade, and teacher representative from second and/or third grades. The SIT made decisions about services using information from SCRF's progress monitoring and outcomes measures, and teacher recommendations.

SCRF school selection.

SCRF schools were selected through an application process, based on need for assistance, and consisted of two cohorts. Cohort 1 schools included schools with participation beginning in year 1, 2004-2005, and Cohort 2 schools included schools with participation beginning in year 4, 2007-2008. Fifteen schools of the original 52 in Cohort 1 were scheduled to rotate out of the SCRF Initiative after the third year because they had participated in a previous state reading initiative. Cohort 2 schools were selected among districts that met eligibility criteria per South Carolina's grant application and were able to recruit an SCRI or SCRF-trained literacy coach. Seven schools were selected in the second application process in 2007-2008, one of which withdrew from the SCRF Initiative in 2009-2010. The number of Cohort 1 and 2 schools, as well as the approximate number of students served at these schools, for each year are summarized in Table 1.

Table 1

SCRf Schools by Academic Year and Cohort

Academic Year	Cohort 1 Schools	Cohort 2 Schools	Approx. # of Students
2004-2005	52	-	13,000
2005-2006	51	-	12,800
2006-2007	48 (+1 new ^a)	-	13,000
2007-2008	30	7	10,000
2008-2009	24	7	8,600
2009-2010	14	6	5,400

Notes. There was one school in each year with K-only, except for 2007-2008 when there were two schools with K-only. In the first three years, one pair of schools consisted of a feeder K-2 school to a 3-5 school.

^aThe one new school that entered the SCRf Initiative in 2007-2008 chose to discontinue in 2008-2009.

Achievement measures.

Multiple reading assessments were administered to students at the SCRf schools. As stated earlier, the *Dominie Reading and Writing Assessment Portfolio* (DeFord, 2000) was used as a progress monitoring tool. The *Dominie* is a diagnostic assessment measuring achievement in reading, writing, spelling, and phonics. The *Stanford Reading First* (Stanford RF) assessment was used as SCRf's measure of achievement for the purposes of program evaluation and was administered twice each year, in the fall and in the spring. In addition, all students in the upper elementary grades completed the state-mandated assessment each spring. Information about the Stanford RF assessment and the state-mandated assessment is presented below.

Assessment for evaluation and reporting. The Stanford RF assessment (Harcourt Assessment, 2004) was used as an outcome measure for program evaluation purposes and for required reporting to the USDOE. Students enrolled in SCRf schools in grades 1-3 completed the Stanford RF assessment in the fall and spring of each school year. The Stanford RF assessment is a version of the norm-referenced, vertically equated Stanford 10. Harcourt

Educational Measurement¹ customized Stanford RF specifically for the national Reading First Initiative.

The total score for a student on the Stanford RF assessment is composed of the score on a multiple choice section and the score on a teacher-administered, oral fluency section. Five components are assessed on the multiple choice section: phonemic awareness, phonics, vocabulary development, reading fluency, and reading comprehension strategies. Two components are assessed on the oral fluency section: speaking vocabulary and oral reading fluency. The total score was used for all analyses presented in this paper.

State-mandated assessment. Data on the performance of students who took the state’s accountability assessment was reported to the U.S. Department of Education annually for all participant schools. The state assessment program begins in grade 3, the highest grade level of SCRF focus. At the start of SCRF implementation, SC’s mandated assessment for state and federal accountability was the Palmetto Achievement Challenge Test (PACT). Student PACT scores were reported in one of four performance levels: Below Basic, Basic, Proficient, and Advanced. For state accountability reporting, “meeting standards” was defined as scoring at or above the Basic level. Beginning in the 2008-2009 school year, SC transitioned to the Palmetto Assessment of State Standards (PASS) as its accountability assessment. The PASS reported scores in three categories in relation to meeting state standards of Not Met, Met, and Exemplary, where “meeting standards” for state accountability was defined as scoring in one of the two top categories.

¹ Harcourt produced the Stanford RF in 2004. Pearson Education, Inc. began producing this assessment later.

Summaries of Key SCRF Studies

The sections below highlight key SCRF studies that can inform R2S in SC. Summaries of achievement results over multiple years of the SCRF Initiative are presented for both the Stanford RF assessment and the state-mandated assessments. Achievement results that highlight patterns of summer loss are also provided. Implementation measurement was documented through the development of an implementation rubric measuring core program components. In addition, the association between implementation and achievement gains was investigated. Intervention strategies were analyzed using both descriptive and inferential analyses of achievement data.

State-mandated Assessment Results

Figure 1 presents the results on the state accountability assessments in reading across all schools that participated in each year of the initiative as well as for a baseline year prior to implementation from the 52 original schools. The percentage of third grade students meeting standards on the PACT assessment showed small increases across all years of the SCRF Initiative with an overall increase of 4.5% from the baseline to the last year of SCRF. SC's assessment changed from PACT to PASS in 2008-2009, thus comparisons from the previous year are not valid. The percentage of students in SCRF schools meeting standards on PASS increased by 5.5% from 2008-2009 to 2009-2010.

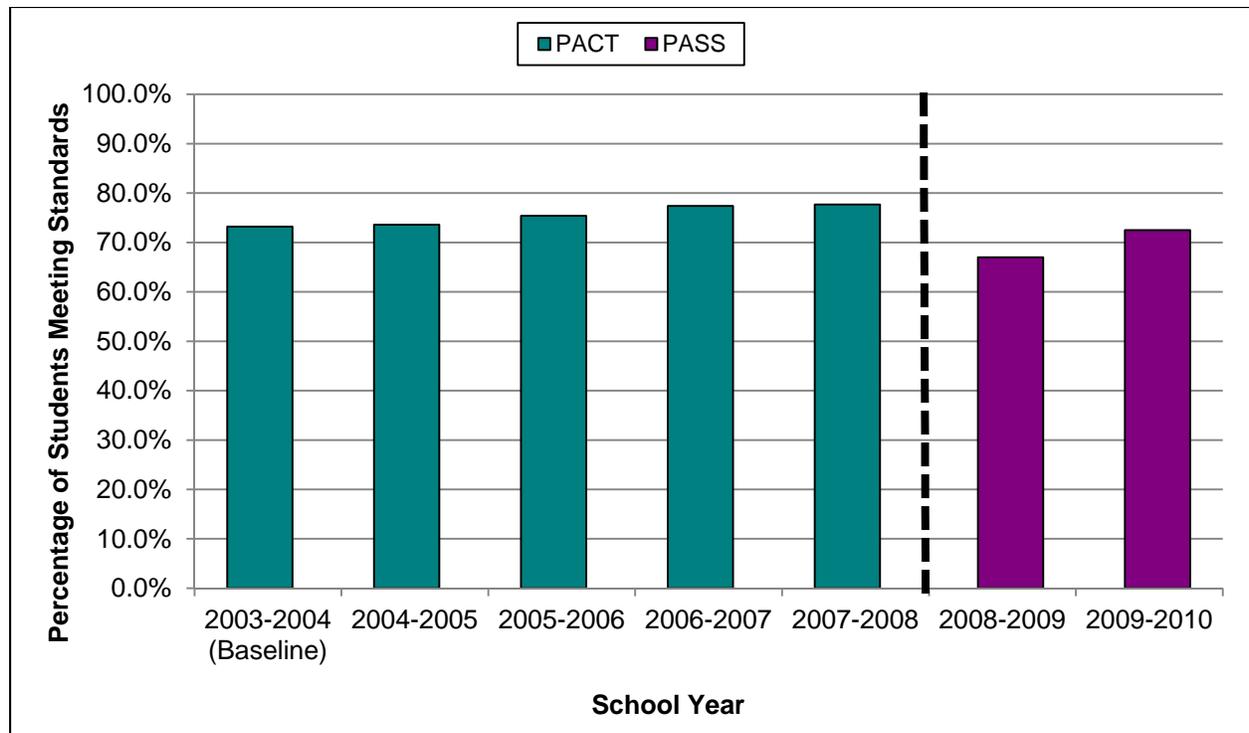


Figure 1. Percentage of third grade students meeting standards on SC state-mandated assessment across all years of SCRF implementation in all SCRF schools.

Stanford RF Achievement Results

Results over all six years of SCRF implementation provide information on the overall trends in achievement² over time. Overall trends are presented for all schools that at any point participated in the SCRF Initiative over the six years of SCRF implementation. Performance level results on the total Stanford RF score are provided for fall and spring across all six years for all students tested. The matched student performance is summarized for schools that participated in consecutive years of SCRF and provides results on mean normal curve equivalent scores for longitudinally matched samples of students.

² Achievement is reported for grades 1-3. One school was an early childhood center with kindergarten participation.

Overall trends for all SCRF schools.

Figure 2 shows the percentage of students who scored At Grade Level³ (AGL) on the fall and spring administrations from 2004 to 2010 for all grades combined. As seen in Figure 1, the percentage of students who scored AGL increased from the fall to spring administration during every year of SCRF. For fall administrations, the percentage of students who scored AGL increased by 12.5% from the first to the sixth year. Considering spring scores, the percentage of students who scored AGL increased by 10.6% from the first to the sixth year. The average increase from fall 2004 to spring 2010 was 19.3%. The average increase in grade 1 students scoring AGL from fall 2004 to spring 2010 was 32.6%, and the increase in grade 3 students scoring AGL across years was 20.2% on average. The percentage of grade 2 students scoring AGL increased by 4.4%, on average, from fall 2004 to spring 2010.

³ At Grade Level on the Stanford Reading First assessment is defined as scoring at or above the 40th percentile.

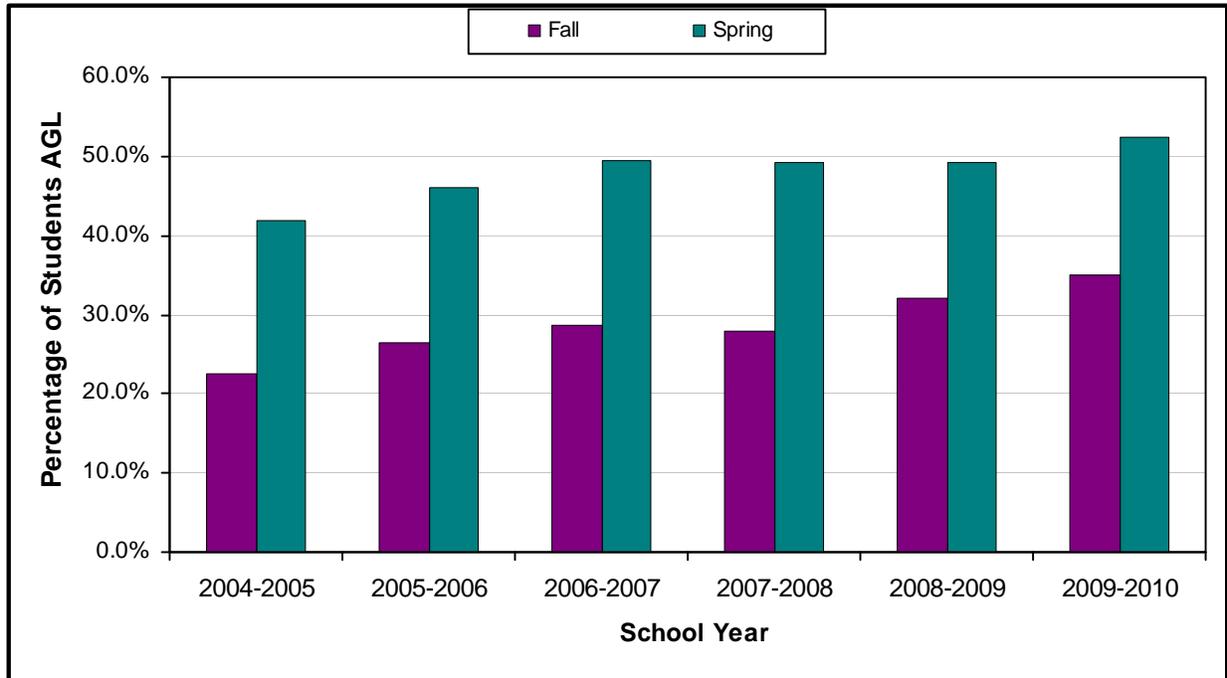


Figure 2. Percentage of students who scored At Grade Level on fall and spring test administrations for all grades combined in all SCRF schools.

Matched sample analysis.

In addition to analysis of absolute performance levels (i.e., proficiency levels), an analysis of normal curve equivalent (NCE) scores on the Stanford RF assessment was conducted after matching individual students who completed consecutive test administrations of the Stanford RF achievement test. The matched sample analysis provides a meaningful measure of growth in reading achievement for students who participated in the SCRF initiative for three consecutive school years in relation to the normative group. If students make similar progress as the norm group, the NCE scores will remain stable over time.

Four groups of students completed six consecutive test administrations of the Stanford RF achievement test. Each matched sample of students participated in SCRF as first graders in the first academic year, second graders in the second academic year, and third graders in the third academic year. For each sample, student scores were matched across the years only if the

students participated in SCRF throughout all three years of implementation and were promoted to the next grade level each year. The first matched sample (2004-2007) included students from 47 schools, the second matched sample (2005-2008) included students from 28 schools, the third matched sample (2006-2009) included students from 23 schools, and the fourth matched sample (2007-2010) included students from 13 schools.

NCE scores were reported based on the total test score on the Stanford RF achievement test. Figure 3 provides an overview of achievement growth over the academic and summer periods for all four SCRF matched samples combined (n=4,486). The cut scores between the Needs Substantial Intervention⁴ (NSI), Needs Additional Intervention⁵ (NAI), and At Grade Level⁶ (AGL) categories on the Stanford RF assessment are indicated by dotted lines.

Results indicated that students showed increases in reading achievement during the school year and substantial losses in achievement over the summer months. As Figure 3 shows, the pattern of achievement growth over academic and summer periods for students in the SCRF program was similar across grade levels: NCE scores increased from fall to spring, and decreased after the summer break. Also, Figure 3 shows that the greatest gain in academic achievement occurred in the first grade, when the average NCE score increased from the NAI category to the AGL category, surpassing the 50th percentile representing average achievement. Although students made less progress in the second grade, the achievement growth in the third grade was substantial, and the mean NCE score of the entire sample at the end of the third grade was within the AGL achievement category. Compared to the norm group, SCRF students made

⁴ Needs Substantial Intervention on the Stanford Reading First assessment is defined as scoring below the 20th percentile.

⁵ Needs Additional Intervention on the Stanford Reading First assessment is defined as scoring between the 20th and 40th percentiles.

⁶ At Grade Level on the Stanford Reading First assessment is defined as scoring at or above the 40th percentile.

relatively larger gains in grades 1 and 3, similar gains in grade 2, and greater loss in the summer.

This pattern of summer loss led SCDE personnel to seek ways to address this problem.

Additional observations of the summer loss pattern are presented in the next section.

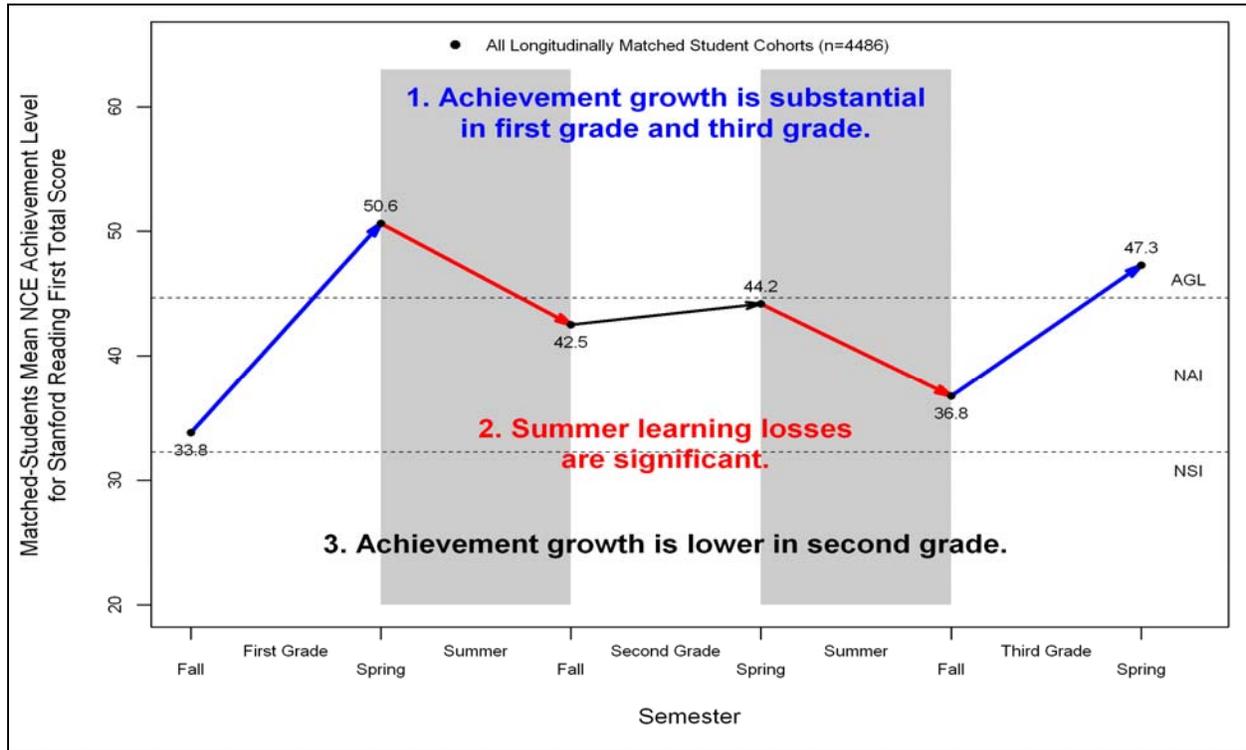


Figure 3. Reading achievement patterns for longitudinally matched SCRF students.

Summer Loss

As seen in the previous section, a prevailing pattern that was observed in the Stanford RF achievement data was increases during the academic year, followed by losses over the summer.

Further analysis by individual and various combinations of demographic subgroups (socio-economic status, race/ethnicity, gender) as well as by initial achievement levels indicates that

this pattern holds across all subgroups observed (Dickenson et al., 2011). Figures 4 and 5

display patterns by combinations of demographics and by quintiles in the fall of grade 1 (initial

achievement) for four longitudinally matched samples of students enrolled in SCRF schools in grades 1, 2, and 3.

The pattern observed for SCRF students is consistent with literature on summer learning. A synthesis of research literature on summer school effects (Cooper, Charlton, Valentine, & Muhlenbruck, 2000) found that students who completed summer remedial work scored higher on standardized assessments than comparison students who did not attend summer school. They further found that programs that provided small-group or individual instruction produced the largest increases in student learning. An earlier review of 39 studies (Cooper, Nye, Charlton, Lindsay, & Greathouse, 1996) found that summer learning loss was equivalent, on average, to 1 month of instruction. Researchers also suggested that the summer period may have the most negative effects on achievement during the early grades as growth in achievement slows over time (Alexander, Entwistle, & Olson, 2001).

Researchers interested in summer learning loss among various groups of students have investigated differential summer learning by characteristics such as student ethnicity, gender, socio-economic status (SES), and prior academic achievement. Results indicate that summer learning loss may have a stronger relationship with students' SES than with other demographic characteristics; this relationship is supported in numerous studies (e.g., Alexander, et al., 2001; Burkham, Ready, Lee, & LoGerfo, 2004; Cooper et al., 1996).

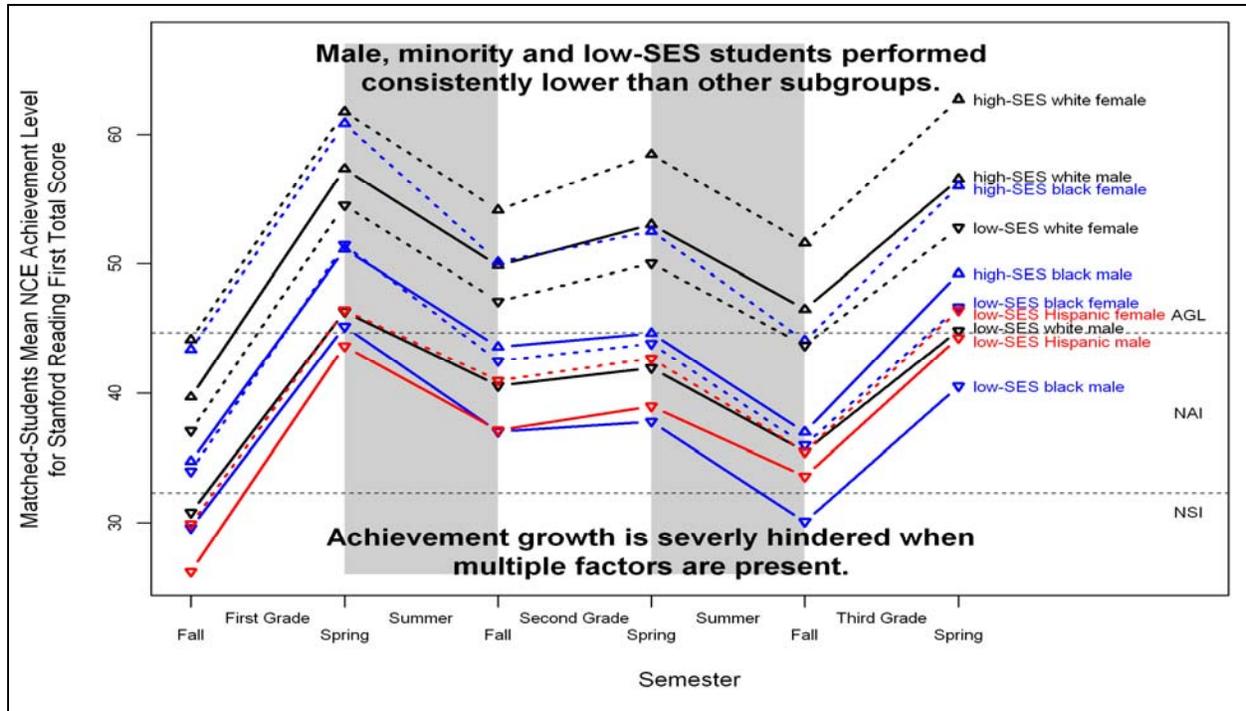


Figure 4. Reading achievement patterns for SCRF students disaggregated by gender, ethnicity, and socio-economic status.

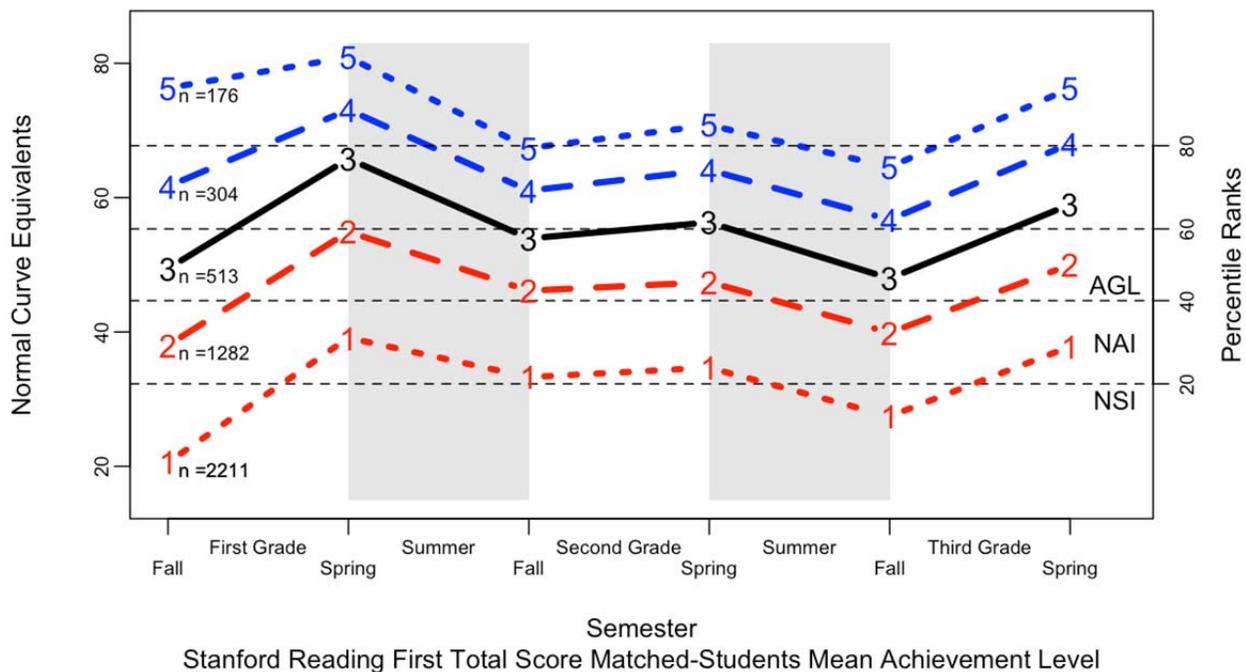


Figure 5. Reading achievement patterns for SCRF students by initial performance quintile in the fall of grade 1.

Implementation Measurement

Measuring implementation fidelity is an often overlooked, but important part of a well-organized program (Field, 1985). According to Century, Rudnick, and Freeman (2010), it is not acceptable to merely measure outcomes to determine if the intervention is fully effective. They urge evaluators to focus on the “why, how, and under what conditions” programs work (p. 30). Throughout the literature, several reasons have emerged for the importance of measuring implementation fidelity in addition to outcomes. First, several studies mentioned the importance of using implementation evaluations to determine whether or not a program is mature enough to merit an outcome evaluation (Leithwood & Montgomery, 1980; Lipsey & Cordray, 2000). Second, measuring implementation fidelity has been used to clarify if the reason for unsuccessful outcomes is because of a flawed program or failed implementation of the program (Mowbray, Holter, Teague, & Bybee, 2003). Third, studies of implementation have been found to assist evaluators with interpreting learning outcomes (Fullan & Pomfret, 1977).

For the SCRF Initiative, an implementation rubric was created to assess fidelity of implementation on various core program components. OPE and SCEPC evaluators collaborated with SCDE personnel to develop an instrument to measure SCRF implementation for both formative and summative purposes. The rubric components were identified by SCRF personnel as core elements to ensure a successful school literacy program based on their experience with successful mature programs. The evaluation team, with expertise in measurement, assisted with translating how program personnel described various implementation levels of each component into descriptors for a rating scale.

The SCRF implementation rubric was initially designed in fall 2008, in part, to assess the level of program implementation that Cohort 1 schools and districts had achieved during the five

years of the initiative. SCRF staff used these data, in conjunction with data on achievement gains, to make decisions about the continuation of specific SCRF schools for the 2009-2010 school year. The rubric was also administered in spring 2010 to six Cohort 2 schools that participated in SCRF for only the most recent three years (2007-2008 through 2009-2010) in order to provide summative data for process evaluation. The rubric was shared with these six schools in 2008-2009 as a shared accountability tool and for formative evaluation purposes. Additional information about the development of the implementation rubric is provided in the 2008-2009 annual SCRF evaluation report (Gareau et al., 2009).

Implementation rubric components.

The implementation rubric contained eight main components that stakeholders identified as essential to the initiative based on their experience with earlier reading initiatives and the early years of SCRF. Three of the main components included sub-components. Table 2 identifies each essential SCRF component, sub-components (where applicable), and the number of items used to measure each component. As noted in the right-hand column, certain items were completed by all participant groups, while other items were completed by SCDE personnel only.

Individual items on the rubric employed one of three response scales including:

(1) Dichotomous (Yes, No), (2) Likert scale (ranging from 0 [*a*] to 3 [*d*], with 0 [*a*] representing the lowest level of implementation and 3 [*d*] representing the highest level of implementation for an individual item⁷), and (3) List (response of *a*, *b*, *c*, or *d* was determined by the number of items chosen from a list). See Figure 6 for examples of each measurement type.

⁷ Two items within the *Assessment* component had a range of 0 [*a*] to 2 [*c*].

Table 2

Essential SCRF Components, Sub-components, and Number of Items

Essential SCRF component	Sub-component(s)	Number of items
Leadership	- School Leadership Team (SLT) - Scheduling - Principal Observations	14 (3 out of 14 completed by SCDE only)
Assessment	- Stanford for Outcome-based - Dominie for Screening/Progress Monitoring/Diagnostic	9 (1 out of 9 completed by SCDE only)
School Intervention Team	N/A	5 (1 out of 5 completed by SCDE only)
Intervention strategies	N/A	4
Instruction	N/A	8
Professional development	- Study Group - School Literacy Coach	12 (3 out of 12 completed by SCDE only)
Environment	N/A	6
District	N/A	2

Implementation rubric scoring.

The implementation rubric was completed by five personnel for each school: two individuals at the school level and three individuals at the state level. At the school level, the school's principal and project director completed the rubric. At the state level, the regional intervention specialist and regional literacy specialist responsible for the school completed the rubric, as well as a SCDE liaison.

RESEARCH AND EVALUATION FROM SOUTH CAROLINA READING FIRST

	Response				
Example: Dichotomous	<input type="checkbox"/> Yes <input type="checkbox"/> No	Study group sessions meet at least twice per month after school for a minimum of 45 contact hours.			
	Response	a	b	c	d
Example: Likert Scale		Required participants do not arrive on time and are not present for the duration of study group meetings.	Some required participants arrive on time and are present for the duration of study group meetings.	Most required participants arrive on time and are present for the duration of study group meetings.	All required participants arrive on time and are present for the duration of study group meetings.
	Response	a	b	c	d
Example: List		The literacy coach at this school demonstrates 0-2 of the following characteristics.	The literacy coach at this school demonstrates 3-4 of the following characteristics.	The literacy coach at this school consistently demonstrates 5-6 of the following characteristics.	The literacy coach at this school consistently demonstrates 7-8 of the following characteristics.
<p>Please check all that apply to their practices at this school. <i>The Literacy Coach at this school...</i></p> <input type="checkbox"/> plans and facilitates on-going, <i>responsive</i> study groups. <input type="checkbox"/> helps to connect the theory behind teacher practices in coaching and study group settings. <input type="checkbox"/> encourages teachers to use active listening. <input type="checkbox"/> encourages teachers to be active kidwatchers. <input type="checkbox"/> encourages teachers in reflective practice to improve instruction. <input type="checkbox"/> encourages teachers to try new instructional practices to improve student engagement and learning. <input type="checkbox"/> assists teachers in administering, scoring, recording, sharing, analyzing, and interpreting student data to inform their instruction. <input type="checkbox"/> assists in the identification and implementation of interventions.					

Figure 6. Examples of item types used on the SCRF implementation rubric.

Implementation and achievement association.

A study by Ishikawa and colleagues (2010) analyzed the relationship between student achievement gains and fidelity of implementation utilizing data from 18 SCRF schools. Student achievement gains on the Stanford RF assessment total score were examined by differences in NCE scores between the spring 2009 value minus the fall 2004 value at the school level. This provided an achievement gain measure appropriate to the five years implementation at the time⁸. For implementation rubric scores, nine possible implementation rubric aggregations were examined: the total implementation rubric score and its eight component sub-scores. Scores for the five raters for each school were averaged yielding a mean total implementation rubric score for each school.

To establish a summative assessment of program effectiveness, regression and correlation analyses were conducted at the school level to test whether schools with a higher fidelity of implementation were associated with greater achievement gain outcomes. For formative assessment purposes, correlation analyses were conducted at the school level to explore which implementation rubric components had the highest degree of association with student gains on the Stanford RF assessment. This process was intended to target improvement efforts in further development of the implementation rubric instrument itself for the 2009-2010 summative administration.

Regression analysis. A regression of Stanford RF total reading achievement gains on the mean total implementation score at the school level revealed a statistically significant positive relationship. On average, schools with higher mean total implementation scores had higher

⁸ This study was conducted prior to the final year of SCRF when data through spring 2009 were available.

Stanford RF achievement gains computed as a difference in all-student NCE averages from the first fall semester to the last spring semester (spring 2009 minus fall 2004). As shown in Figure 7, with a slope coefficient of $b = 0.358$ (and a one-sided alternative p-value of 0.0217), the relationship was moderately positive and statistically significant.

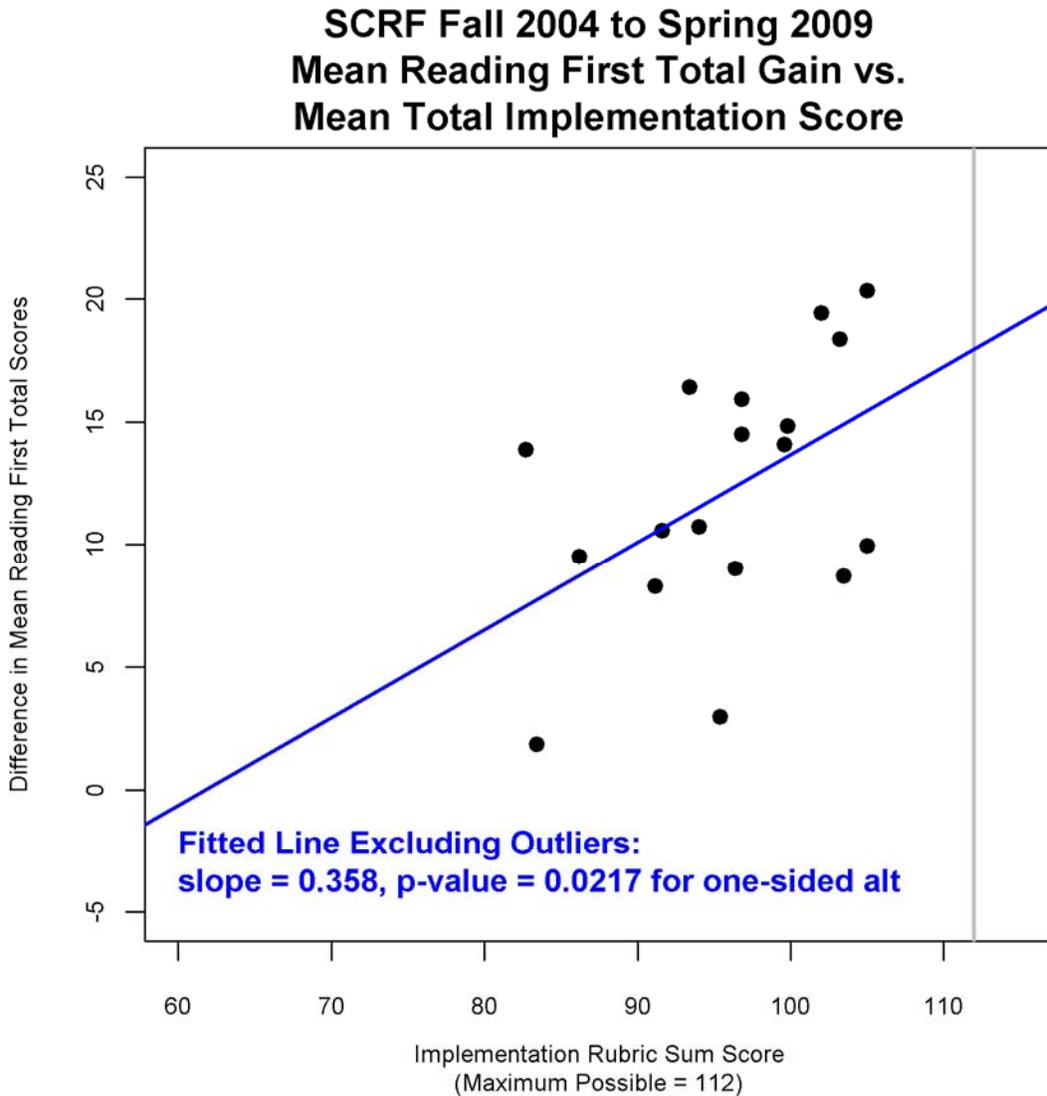


Figure 7: Achievement outcome gain versus mean total implementation rubric score.

Correlation analysis. For formative assessment purposes of rubric development, the relationship between Stanford RF achievement scores and eight implementation rubric component scores using correlation analysis was investigated. A Pearson product-moment correlation was calculated between the NCE achievement gain and the total rubric score, as well as with each of the eight implementation rubric component scores. Results from the correlation analysis are presented in Table 3. For achievement gains computed using the total Stanford RF score, correlations with three components (Instruction, Literacy Environment, and District Support) were significant at the .05 level with correlations near .50. Intervention Strategies also showed statistical significance at the .10 level, albeit at a lower effect size ($r=.37$, $p\text{-value}=.068$).

Table 3

Correlations between Stanford RF Fall 2004 to Spring 2009 Change and Implementation Rubric Components

Achievement	Implementation	N	<i>r</i>	<i>p</i>
Reading First Total	Total Implementation Score	18	.48	.0217
Reading First Total	Component Score 1: Leadership	18	.25	.1547
Reading First Total	Component Score 2: Assessment	18	.28	.1296
Reading First Total	Component Score 3: School Intervention Team	18	-.03	.5407
Reading First Total	Component Score 4: Intervention Strategies	18	.37	.0680
Reading First Total	Component Score 5: Instruction	18	.51	.0146
Reading First Total	Component Score 6: Professional Development	18	.07	.3889
Reading First Total	Component Score 7: Literacy Environment	18	.49	.0186
Reading First Total	Component Score 8: District Support	18	.40	.0498

Notes. Bold font indicates statistical significance at the .05 level and italicized font indicates statistical significance at the .10 level.

Intervention Studies

The authors conducted a descriptive analysis of achievement change over each school year for students who received intervention and those who did not. In addition, Dickenson and Young (2010) conducted an inferential study using a hierarchical linear model (HLM) for growth to examine longitudinal achievement data for students who participated in SCRF for three consecutive years. Summaries and results of these analyses are provided below.

Descriptive analysis across years.

The reading scores of students who received intervention and those who did not were compared from fall to spring each academic year by grade level. Looking year by year for each grade level, the percentage of students whose reading score improved from fall to spring was higher for those who received intervention than for those who did not in 10 of the 18 analyses conducted (see Table 4). In the first two years, students who received intervention improved performance categories (e.g., moved from NAI to AGL) at all three grade levels at a lower rate than students who did not receive intervention. In years 5 and 6, these trends reversed with the greatest improvement seen in grade 1, which had the highest frequency of students served in intervention. Students who received intervention in grade 1 improved performance categories by 9.6 percentage points in year 5 and by 8.4 percentage points in year 6 more than those who did not receive services. This might reflect the time needed to properly train interventionists on strategies and to fully implement those strategies to assist struggling readers.

Table 4

Frequency of Additional Intervention and Percentages of Reading Performance Improvement by Grade and School Year

School year	Grade	Number of students		Percentage of students who improved at least one performance category		Difference ^a
		No Intervention	Intervention	No Intervention	Intervention	
2004-2005	1	2,251	1,011	60.7	55.4	-5.3
	2	2,604	405	26.8	22.5	-4.3
	3	2,712	202	45.4	44.1	-1.3
2005-2006	1	2,071	1,117	59.8	55.2	-4.6
	2	2,296	716	24.7	23.7	-1.0
	3	2,196	558	45.4	43.2	-2.2
2006-2007	1	2,188	1,000	56.2	58.4	2.2
	2	2,170	693	27.5	27.6	0.1
	3	2,274	557	46.0	44.7	-1.3
2007-2008	1	1,479	916	58.8	61.0	2.2
	2	1,758	585	26.2	32.0	5.8
	3	1,802	452	49.1	43.8	-5.3
2008-2009	1	1,355	728	51.1	60.7	9.6
	2	1,589	508	22.3	23.6	1.3
	3	1,604	442	42.8	48.0	5.2
2009-2010	1	902	501	49.3	57.7	8.4
	2	947	322	18.0	20.2	2.2
	3	1,103	225	38.0	41.3	3.3

^aPositive numbers indicate higher percentage of improvement among students who received additional intervention.

Study of intervention growth.

In addition, Dickenson and Young (2010) conducted an inferential study that sought to determine the effectiveness of SCRF intervention methods at accelerating growth in reading achievement. Intervention was provided to students based on low performance on measures of basic reading skills. Because inclusion for services was based on academic need, an equivalent

comparison group was not available. To determine intervention effectiveness, growth rates were compared between students who received intervention and those who did not at each grade level.

An HLM growth model was used to investigate the statistical significance of achievement growth for students who received intervention services compared to those who were not served. A three-level HLM (e.g., Raudenbush & Bryk, 2002), accounting for the nested structure of multiple tests within students and students within schools, was used to model longitudinal growth. The model included linear and quadratic growth parameters to account for change in the growth rate over time. If the intervention is effective, then achievement growth will accelerate at a faster rate for students who receive intervention services.

Three longitudinally matched samples of students were utilized in this study from schools that used SCRF interventions focused on Reading Recovery® and small group interventions. The students in the matched samples were enrolled in SCRF schools for three consecutive years and progressed from grade 1 to grade 2 to grade 3. The 2004-2007 matched sample included 1,500 students in 43 schools, the 2005-2008 sample included 927 students in 25 schools, and the 2006-2009 sample included 881 students in 21 schools.

Students were identified for intervention services based on academic need and available intervention resources at the school. Approximately 40% of students in the schools were served in intervention in each sample at some point over the three years. Of these, approximately 22% received intervention in just one grade. Those who required intervention in multiple grades had persistent challenges with reading.

Summary of results. Grade 1 students who received intervention tended to make greater gains on the Stanford RF assessment initially than those who did not receive intervention, which leveled off in later administrations. This provides evidence of the intervention accelerating

growth for first grade students to catch up with their peers. The rate of growth for students who received intervention in grade 2 tended to accelerate over time, thus providing evidence of effectiveness of helping these students to catch up at this grade level. For students served in grade 3, it is difficult to judge the success of the intervention in accelerating growth as the data concludes with grade 3. Students being served in grade 3 have either had persistent difficulty with reading or were newly identified as needing additional support. Results for intervention in grade 3 were inconclusive. Additional time is needed to determine whether intervention will be effective for students served in grade 3. The findings support the effectiveness of SCRF intervention services toward closing the achievement gap in early grades.

Discussion of Lessons Learned

The research and evaluation studies associated with the SCRF Initiative contribute to several “lessons learned” which will be valuable in the implementation of R2S. These lessons address the importance of collaborative evaluation, the commitment and involvement of all stakeholders, the critical role of assessment, summer loss, implementation fidelity, and program outcomes and expectations.

Importance of Collaborative Evaluation

The evaluation of the SCRF Initiative was a collaborative effort between external evaluators and personnel at the SCDE. The evaluation team participated in the design of the initiative during the development of the proposal for federal funding, assisted in the identification of appropriate assessment instruments, developed participant surveys and implementation rubrics, and conducted ongoing analyses to inform continual project improvement. This collaboration involved regular meetings in which project implementation and evaluation activities were planned and evaluation results (e.g., surveys, student achievement)

were shared. The evaluation team provided the collaborative with the expertise required to examine trends in achievement scores using matched samples, longitudinal analyses, and hierarchal linear model analyses. A similar collaborative design will benefit R2S in providing data to review progress in student achievement and monitor implementation for continued program improvement.

Commitment and Involvement of All Stakeholders

SCRF was unique among state literacy initiatives in the level of commitment and involvement that it required from project schools and districts. State-level SCRF personnel at the SCDE, including the regional specialists, developed very specific memorandums of understanding with the districts which detailed SCRF program requirements including types of curricula and assessments which had to align with SCRF. For example, district administrators had to limit conflicting initiatives at SCRF schools, remove or restrict other testing or programs, support the coaching role as collaborative, not evaluative, etc. Principals had similar responsibilities at the school level and were required to be full participants as learners in the mandatory weekly study groups for teachers conducted by school literacy coaches. Principals also protected the 120 minute uninterrupted reading block, the full-time role of the coach, and other critical elements of the SCRF. School Leadership Teams included the principal, coach, district contact, teacher representative, special education teacher, and media specialist. These teams were responsible for creating the support structures necessary to implement the SCRF Initiative as designed at their schools. School Intervention Teams focused attention on the neediest readers at each school.

The design of the SCRF Initiative, from the required personnel and organizational structures to the comprehensive, ongoing training and support from SCDE staff fostered the

involvement and commitment of all stakeholders needed to build a strong literacy program. In the shared learning of the study groups and training sessions, a common understanding of SCRF and its components was built.

Critical Role of Assessment

All SCRF schools used the same assessment instruments to monitor student progress and achievement, rather than having a variety of instruments in use from school to school. For the first time in the many years of state literacy initiatives, this uniformity of measurement allowed the evaluation team to reliably assess student progress and relate student growth to a variety of other variables such as implementation components. The trends of SCRF gains in student achievement on the Stanford Reading First assessment, a norm-referenced, vertically equated test capable of measuring yearly student growth, were evident (see Figures 2 and 3) and could be assessed reliably across all project schools. For state policymakers to monitor trends in student achievement for R2S, uniform, vertically equated or vertically moderated state-level achievement tests should also be selected for R2S.

In addition, the fall and spring administration of the Stanford Reading First assessment permitted an accurate measurement of student growth during the school year and revealed student achievement declines over the summer. If the SCRF Initiative had been solely measured with a spring test administration, the initiative would have deemed another reform failure as were the majority of other states' Reading First programs with only spring testing. Multiple testing points allowed all stakeholders, and especially the students' coaches and teachers, to see the actual achievement gains of the students during the school year. With the increasing weight placed on teacher evaluation, it is especially crucial to have accurate measurements of student growth that can be attributed to the efforts of state and school personnel.

Summer Loss

Students participating in the SCRF Initiative showed increases in student achievement during the school year and substantial losses in achievement over the summer months. When following students who participated in SCRF for three consecutive years, each year of instruction was followed by a substantial summer learning loss. At the beginning of second grade, the average normal curve equivalent (NCE) score was 8.1 lower than the average score recorded at the end of first grade. Similarly, there was a discrepancy of 7.4 between the mean NCE scores obtained at the end of the second grade and the beginning of third grade. All subgroups of students experienced the same general pattern of school-year gains and summer losses, but substantial differences in achievement scores occurred between subgroups of students. Males, minority students, and students of low socioeconomic status performed consistently lower than other subgroups of students.

Future reading programs should incorporate research-based strategies or programs to mitigate summer loss in student reading achievement, particularly for at-risk students. A recent meta-analysis on summer reading programs, including both classroom and home interventions, identifies effective practices (Kim & Quinn, 2013). As recommended for school-year literacy programs, common assessments should be employed and implementation rubrics used so that elements of successful programs can be identified and disseminated. Many of South Carolina's students could attain higher levels of reading performance if they did not have to overcome significant summer achievement losses each year.

Implementation Fidelity

Fidelity of implementation is critical in understanding how and why programs work (Century et al., 2010). Specification of the critical program components in a tool to measure

implementation fidelity also ensures that all stakeholders have a shared understanding of important program elements and the timeline for implementation. The evaluators, with expertise in measurement, worked collaboratively with SCDE program staff to develop the implementation rubric after the fourth year of SCRF implementation. The rubric components were identified by SCRF personnel as core elements to ensure a successful school literacy program based on their experience with mature programs. The evaluation team assisted with translating how program personnel described various implementation levels of each component into descriptors for a rating scale.

In the SCRF Initiative, the ratings on the implementation rubric used to review school implementation of the program were associated with student achievement. On average, schools with a higher fidelity of SCRF program implementation showed greater student achievement outcome gains. Student achievement gains significantly correlated with the total implementation score (.48), elements of classroom instruction (.51), the school literacy environment (.49), and school district supports (.40). However, some component scores showed no correlation with achievement (e.g., school intervention team, professional development), possibly because of low variability in these measurements across schools or needed refinement of rubric items.

Continued development and use of such instruments in R2S will provide feedback for reviewing the implementation of program components. By measuring not only student achievement outcomes, but also how faithfully the program is implemented compared to its design, the validity of the program evaluation process is increased. This enhanced program assessment affords an opportunity to discover how or why the program succeeded or failed beyond the mere result of ultimate success or failure. Utilizing implementation measurement tools at the outset of

R2S can create a shared understanding of accountability between state-level leadership and districts and schools.

Program Outcomes and Expectations

When new programs are implemented, there is often an expectation of immediate results. SCRF provides a lesson in why expectations should be viewed on a more long term basis. Particularly for programs requiring extensive professional development and changes in routine practices, time is needed to complete the requisite training and build the capacity and skills of program personnel to deliver program services with fidelity.

An example is a delay in immediate results when observing data from the SCRF intervention model. SCRF utilized Reading Recovery® and a similar approach for small groups as its primary intervention model. Reading Recovery® is an evidence-based approach that is recognized for its effectiveness by the What Works Clearinghouse (U.S. Department of Education, 2013). Intervention was new to the state and SCRF schools were among the first to implement a Response to Intervention approach that is now common among SC schools. In the first two years of SCRF, students who received intervention services improved performance categories at all three grade levels at a lower rate than students who did not receive intervention. It appeared as if students receiving extra assistance were actually falling further behind when compared to students who received no help. However, in years 5 and 6, these trends reversed with the greatest improvement seen in grade 1. Further, longitudinal analysis over three years indicated that participation in SCRF interventions significantly accelerated students' growth in reading achievement when provided in early grades. SCRF interventionists received training as they simultaneously implemented new strategies. The focus was on one-to-one strategies in year 1 and small group strategies in year 2. It was not until the third year that interventionists were

fully trained in both types of intervention methods and positive achievement results began to emerge. Thus, R2S programs might benefit from formative evaluative feedback about their programs in the initial years and summative reviews should be initiated later in the implementation of the program.

References

- Alexander, K. L., Entwisle, D., & Olson, L. S. (2001). Schools, achievement, and inequality: A seasonal perspective. In G. Bohrman & M. Boulay (Eds.), *Summer learning: Research, policies, and programs* (pp. 25-52). Mahwah, NJ: Erlbaum.
- Bennett, H.L., Dickenson, T.S., Monrad, D.M., Gareau, S., Johnson, R.L., Gilmore, J., Ishikawa, T., Miller, K.M., Mindrila, D., & Morgan, G. (2010, November). *South Carolina Reading First Initiative Evaluation Report 2009-2010*. Columbia: University of South Carolina, Office of Program Evaluation and the South Carolina Educational Policy Center.
- Bennett, H., Gareau, S., Dickenson, T., Monrad, D., Johnson, R., Galande, M., Mindrila, D., Ermer, B., Gilmore, J., Smith, J., Rawls, A., Gay, J., & Sesso-Dahlke, K. (2007). *South Carolina Reading First Initiative Evaluation Report 2006-2007*. Columbia, SC: University of South Carolina, Office of Program Evaluation and South Carolina Educational Policy Center.
- Bennett, H., Gareau, S., Dickenson, T., Monrad, D., Johnson, R., Petrusis, R., Gilmore, J., Smith, J., Rawls, A., Galande, M., Chiuzan, C., Gay, J., Gambone, L., Price, K., & Mindrila, D. (2008, November). *South Carolina Reading First Initiative Evaluation Report 2007-2008*. Columbia, SC: University of South Carolina, Office of Program Evaluation and South Carolina Educational Policy Center.
- Burkham, D. T., Ready, D. D., Lee, V. L., & LoGerfo, L. F. (2004). Social-class differences in summer learning between kindergarten and first grade: Model specification and estimation. *Sociology of Education*, 77(1), 1-31.
- Century, J., Rudnick, M., & Freeman, C. (2010). A framework for measuring fidelity of implementation: A foundation for shared language and accumulation of knowledge. *American Journal of Evaluation*, 31(2), 199-218.
- Cooper, H., Charlton, K., Valentine, J.C., & Muhlenbruck, L. (2000). Making the most of summer school: A meta-analytic and narrative review. *Monographs of the Society for Research in Child Development*, 65, 1-117.
- Cooper H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, 66, 227-268.
- DeFord, D. (2000). *Dominie Reading Writing Assessment Portfolio, Reading Assessment: Teachers Guide Part I*. Carlsbad, CA: Dominie Press, Inc.
- Dickenson, T.S. (2008, March). *Three Year Growth in Achievement for South Carolina Reading First*. Presented at the annual meeting of the American Educational Research Association, New York, NY.
- Dickenson, T.S. (2009, November). *Using Hierarchical Linear Growth Models to Investigate Reading Achievement Gaps*. Paper presented at the annual meeting of the American Evaluation Association, Orlando, FL.

- Dickenson, T.S., McGuiness, P.L., Monrad, D.M., Johnson, R.L., Gareau, S.J., Gilmore, J.A., McCorkendale, S.C., & Payne, J.R. (2005, November). *South Carolina Reading First Initiative Highlights of the 2004-2005 Evaluation Results*. Columbia: University of South Carolina, Office of Program Evaluation and the South Carolina Educational Policy Center.
- Dickenson, T.S. & Monrad, D.M. (2007, April). *Growth in Student Achievement for the First Two Years of South Carolina's Reading First Initiative*. Paper presented at the annual meeting of the American Educational Research Association, Chicago IL.
- Dickenson, T.S., Monrad, D.M., Johnson, R.L., & Wills, P. (2008, March). *Evaluation of South Carolina Reading First*. Individual session within symposium titled *Evaluating the Impacts of Reading First: Results from Three States and Policy Considerations*. Presented at the annual meeting of the American Educational Research Association, New York, NY.
- Dickenson, T.S., Monrad, D.M., May, R.J. & Bennett, H.L. (2011, April). *Investigation of Reading Achievement Growth over Academic and Summer Periods*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Dickenson, T.S. & Smith, J. (2009, April). *Factors Related to Growth in Achievement for South Carolina Reading First Schools*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Dickenson, T.S., Smith, J., & Gilmore, J. (2007, April). *Selection of Comparison Schools for South Carolina's Reading First Initiative with Follow-Up HLM Analysis*. Paper presented at the annual meeting of the American Educational Research Association, Chicago IL.
- Dickenson, T. & Young, J. (2010, November). *Evaluation of Reading Intervention Effectiveness Using Growth Models*. Paper presented at the annual meeting of the American Evaluation Association, San Antonio, TX.
- Field, J.O. (1985). Implementing nutrition programs: Lessons from an unheeded literature. *Annual Review of Nutrition*, 5, 143-172.
- Fullan, M., & Pomfret, A. (1977). Research on curriculum and instruction implementation. *Review of Educational Research*, 47(2), 335-397.
- Gareau, S., Bennett, H., Dickenson, T., Monrad, D., Johnson, R., Chiuzan, C., Galande, M., Gilmore, J., Ishikawa, T., Loudermilk, B., Mindrila, D., Price, K., Rawls, A., & Smith, J. (2009, November). *South Carolina Reading First Initiative Evaluation Report 2008-2009*. Columbia, SC: University of South Carolina, Office of Program Evaluation and South Carolina Educational Policy Center.
- Harcourt Assessment, Inc. (2004). *Stanford Reading First: Technical Data Report*. United States: Harcourt Educational Measurement.
- Ishikawa, T., Dickenson, T., Monrad, D., Bennett, H., Mindrila, D., Miller, K.M., and Gareau, S. (2010, November). *A Study of the Relationship between Fidelity of Implementation and Achievement Outcomes*. Paper presented at the annual meeting of the American Evaluation Association, San Antonio, TX.

- Kim & Quinn (2013). The effects of summer reading on low-income children's literacy achievement from kindergarten to grade 8: A meta-analysis of classroom and home interventions. *Review of Educational Research*, 83(3), 386-431.
- Leithwood, K.A., & Montgomery, D.J. (1980). Evaluating program implementation. *Evaluation Review*, 4(2), 193-214.
- Lipsey, M. & Cordray, D. (2000). Evaluation methods for social research. *Annual Review of Psychology*, 51, 345-375.
- Morgan, G.B., Dickenson, T.S., & Young, J.E. (2012, April). *Investigation of Reading Intervention Effectiveness Using Propensity Score Matching*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, British Columbia, Canada.
- Mowbray, C.T., Holter, M.C., Teague, G.B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation*, 24(3), 315-340.
- National Council of Teachers of English. (2008). The South Carolina Reading Initiative: Sponsored by the Governor's Institute of Reading at the State Department of Education, background. National Council of Teachers of English Website. Retrieved online November 9, 2009 from http://www.ncte.org/library/NCTEFiles/Resources/Magazine/SCRI_Report.pdf.
- National Reading Panel (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction [on-line]. Available: <http://www.nichd.nih.gov/publications/nrp/report.cfm>.
- Pearson Learning Group. (2004). *Dominie Writing and Reading Assessment Portfolio*. Retrieved November 8, 2009 from <http://www.pearsonschool.com/index.cfm?locator=PSZ16e&PMDBSUBCATEGORYID=28139&PMDBSITEID=2781&PMDBSUBSOLUTIONID=&PMDBSOLUTIONID=6724&PMDBSUBJECTAREAID=&PMDBCATEGORYID=3289&PMDBProgramID=19382>
- Pinnell, G. S., DeFord, D. E., & Lyons, C. A. (1988). *Reading Recovery: Early intervention for at-risk first graders (Educational Research Service Monograph)*. Arlington, VA: Educational Research Service.
- Raudenbush, S. & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd edition)*. Thousand Oaks, CA: Sage Publications.
- Sesso-Dahlke, K., Dickenson, T., Monrad, D., Johnson, R., Gilmore, J., Rawls, A., Mindrila, D., Smith, J., McGuinness, P., Wills, L., Gay, J., & Ermer, B. (2006). *South Carolina Reading First Initiative Evaluation Report 2005-2006*. Columbia, SC: University of South Carolina, Office of Program Evaluation and South Carolina Educational Policy Center.
- U.S. Department of Education, Institute of Education Sciences, What Works Clearinghouse. (2013, July). *Beginning Reading intervention report: Reading Recovery®*. Retrieved from <http://whatworks.ed.gov>