

]]]
**THE FINAL YEAR OF THE
STEDMAN LITERACY PROJECT:
1999-2000**
]]]

Center for Policy Research

**1570 Emerson Street
Denver, Colorado 80218
303/837-1555**

-] Jessica Pearson, Ph.D.**
-] Nancy Thoennes, Ph.D.**
-] Sarah Birkeland, M.A.**
-] Carolyn M. Lupe, M.A.**

August 2000

] Executive Summary]

The Stedman Literacy Pilot Project was initiated by Summer Scholars in September 1997 to enhance the performance of students during the regular, nine-month school year by exposing them to more intense instructional formats. In 1999-2000, the Pilot served 188 first, second, and third graders, 71 percent of whom were eligible for free or reduced price school lunches, 80 percent of whom were African-American, and only two percent of whom were Anglo. During its three years of operation, the \$100,000 per year Pilot included the following:

-] The creation of a daily two-hour uninterrupted literacy block for all first and second graders in 1997-1998 and a 2.5 hour literacy block for first, second, and third graders in 1998-1999 and 1999-2000;
-] Reductions in the student/teacher ratio during the morning literacy block to 18:1 and, when fully staffed with paraprofessionals and volunteers, the student/staff ratio to 6:1;
-] The assignment of all first, second, and third grade students to relatively homogeneous groups according to reading ability during the daily literacy block;
-] Reductions in pull-outs for ESL classes and other rival activities during the morning literacy classes;
-] The retention of a part-time literacy specialist and the provision of regular staff training on state-of-the-art teaching techniques;
-] The adoption of an effective instructional strategy known as the balanced literacy approach to reading and writing;
-] The organization of two family literacy events at the school to involve parents more fully in their children's education;
-] The conduct of an extracurricular reading project to encourage home reading ("Million Minutes") by offering students prizes for reading.

The evaluation of the Pilot Project involved the administration of reading tests to Stedman students at the start and end of each school year, classroom observations, and focus groups with teachers. In addition, we compared the performance of Stedman students on various individual and standardized reading tests with their counterparts in

DPS and a cross-section of Colorado school districts. We also compared Stedman results with those gleaned in national evaluations of reduced class size and other intensive literacy programs.

In the third and final year of the Stedman Literacy Pilot, as in the first two years, the most notable evaluation finding is that students in smaller classes with intense literacy instruction experience strong and significant gains on individually-administered reading tests. These gains exceed those demonstrated by over 3,000 second and third grade students in regular Colorado classrooms using identical reading tests as part of a Multi-district study. At the end of the school year, two-thirds of first graders, 85 percent of second graders and 98 percent of third graders successfully read passages that are considered to be at or above grade level. Only about half of the students in the Multi-district study demonstrated this level of reading skill.

Like Summer Scholars teachers, Stedman teachers attribute the exceptional gains Stedman students make to their daily participation in intensive literacy classes characterized by individual and small group reading activities with similarly situated children in low student/staff ratio settings.

As in past years, student performance on standardized reading tests was more mixed. While first and second graders performed the best on the ITBS, with the average student performing slightly below the national average (42 and 45 percent), 81 percent of Stedman's third graders scored below grade level, with the average student demonstrating a lag of one year in reading proficiency. These patterns are consistent with those demonstrated by DPS students as a whole and Black students in the District, although it is encouraging that Stedman students registered stronger improvements on the ITBS at every grade while DPS students made more modest improvement at the second grade level and actually fell behind at grades 3-5.

Stedman's performance on the CSAP, Colorado's test to measure student performance relative to a standard, was extremely encouraging, with the proportion of third graders testing at or above proficiency rising from 37 to 48 percent. This proficiency level is slightly higher than the district average of 47 percent and substantially higher than the 39 percent proficiency level for the district's Black third-graders.

Our effort to gauge the cumulative and more lasting effects of the Stedman Literacy Pilot project also yielded mixed results. While those who were in the project for all three years read more proficiently on individually-administered reading tests than those in the project for one or two years, they performed no better on the CSAP or the ITBS. And while Stedman's fourth graders who transitioned into regular-sized classrooms scored worse than District fourth graders as a whole on the ITBS and placed in the 36th percentile (but comparable to Black fourth graders), they registered dramatic gains during the school year and improved by nearly 10 "percentage points," as compared with

a drop of 0.4 among fourth graders in the District as a whole. Stedman's fourth graders also improved on the CSAP, with 35 percent achieving a proficiency rating as compared with 29 percent in 1999. Although they still fell below the district average of 38 percent, they out-performed Black DPS fourth graders, 30 percent of whom scored at or above proficiency on the CSAP in Spring 2000.

No test does a perfect job of measuring what students know or how they perform; it is impossible to reconcile the different pictures of achievement that each test provides. It is clearly more difficult to get students to successfully apply their improved ability to read and comprehend short passages to other material and in less familiar testing formats. By the higher primary grades (3-5), when standardized tests become harder and more independent, fragile readers may be even more disadvantaged than their younger counterparts and score less consistently on different types of tests.

Our assessment is weakened by the lack of a matched group of students in regular classes against whose progress Stedman's students may be compared, the small number of students in the Literacy Pilot, and its implementation in just one school. Stedman had some troubles with staff absenteeism and teacher turnover (especially in the third grade) that may well have reduced effects on achievement.

The Stedman findings of strong improvement in reading skill with few parallel gains on standardized tests are consistent with the mixed results reached in evaluations of other achievement programs for low-income students like the widely-adopted Success for All (SFA). While researchers find that while SFA students improve in reading and outscore their counterparts from comparison schools, only 12.5 percent of SFA students were reading at or near grade level after five years of program participation.

The one approach to achievement for low-income children that evaluators do find to have clear and consistent achievement benefits is being taught in small classes with a teacher: student ratio of 1:15 in Kindergarten through grade 3. Follow-up studies with students in a Tennessee study show that the advantages of small classes are retained after students return to regular-size classes, with the most recent research showing benefits all through high school, although the biggest average gains in performance on standardized tests (4 percentile points) happen in the first year of assignment to a small class.

These studies, along with the Stedman evaluation, underscore the benefits of intensive educational policies for low-income students at the primary level. Even if they do not always lead to grade-level performance on norm-referenced tests, they consistently produce strong and significant gains in reading skill among the most fragile learners and narrow the achievement gap between White students and racial minorities on newer, standards-based performance tests like the CSAP. This is an impressive achievement, especially for a program that served 188 students with an annual budget of only \$100,000

(a cost of \$532 per student per year). There is research evidence that more extensive school reform efforts can have an ever bigger impact. For example, in Tennessee, where small class size has been used the longest and with the largest numbers, students who experienced small classes in K-3 show positive and significant long-term effects.

] Recommendations]

The Stedman Literacy Pilot ended in June 2000; Summer Scholars will cease to play a formal role in fund-raising, program design, implementation, and evaluation. Fortunately, the program has become firmly rooted in the school; the principal and teachers hope to incorporate elements of the Literacy Pilot in future instructional formats on a long-term basis. The following recommendations are designed to help Stedman staff with program continuity and to help guide Summer Scholars in its advocacy on behalf of Denver's disadvantaged readers.

] Collaborations between schools and community groups can be extremely effective ways of generating resources, energy, and focus to support literacy interventions. Through the fund-raising efforts of Summer Scholars, Stedman benefitted from an additional \$100,000 per year, which was used to hire part-time teachers and reading assistants to reduce class size for 2.5 hours each day. Summer Scholars also recruited volunteers to enhance literacy instruction further. Finally, the collaboration led to an enhanced planning and evaluation effort which heightened program visibility and accountability. Although student performance on norm-referenced tests is mixed, the results are extremely impressive on individually-administered reading tests and standards-based performance tests. Schools should actively pursue collaborations with community groups that have expertise in achievement programs for low-income students to augment their efforts to promote student literacy.

] It is recommended that Stedman continue to offer small, homogeneous classes for literacy instruction for at least 2.5 hours per day. All relevant internal and external resources should be marshaled for this effort with the goal of achieving literacy classes with a student/teacher ratio no greater than 15:1. In addition to pursuing relevant federal, state and local grants, Stedman should continue to explore other possible reorganizations such as the use of Title I reading staff, the Assistant Principal, and other non-classroom personnel, to support regrouping for literacy classes with lower student/teacher ratios.

-] It is recommended that Stedman invest its limited resources in teachers rather than reading assistants and other paraprofessionals who have displayed an unacceptable level of absenteeism and turnover. Due to high levels of staff turnover and absenteeism, some literacy classes operated without full staffing. This meant that teachers had to rely more on large-group instruction and could not offer all students a daily opportunity to read appropriate material in a small group with similarly situated students. The Tennessee STAR Project found no achievement benefits to large classes with or without an aide; the only benefits came from exposure to teachers in low ratio settings.
-] Students who attend school regularly make more progress than those who don't and many Stedman students continue to enroll late, leave early and attend irregularly. The loss of the school's Family Resource Coordinator has made lines of responsibility unclear for monitoring attendance and intervening with families. Stedman must implement an aggressive attendance program and identify personnel who will quickly flag children who are repeatedly absent or tardy. Since absenteeism is significantly greater among children in single versus two parent homes, this might also be a population to target for intervention by DPS social workers.
-] Whether they begin the year reading at or below grade level, students who engage in extracurricular reading do better on year-end tests. It is recommended that teachers actively support the Million Minutes Project which provides incentives for home reading.
-] While students in the lowest grades make the most gains and intervention should appropriately focus on students in grades K-3, there is a definite need for additional remediation at later grades. Before and after school programs should include one-on-one tutoring, and small group support in reading. Older students also need summer school interventions to prevent the dramatic loss in reading skill that typically occurs for low-income students during the summer months.
-] Three years is not a long enough test for a pilot project aimed at demonstrating the effects of an intensive literacy program. Nor is implementation in a single school setting reliable. Complex programs take time to get right. And since there is always a range in implementation depending on school culture and climate, any test of a new program should be done in multiple settings in order to assess the impact of high and weak program implementations.
-] The Stedman Literacy Pilot Project was implemented with a minimal amount of advance planning and attention to school and community "buy-in." Other literacy initiatives in the Denver area are dedicating a full year to the planning process. Summer Scholars should monitor these newer efforts and assess the appropriate mix of planning and action before undertaking another school year program.

] Summer Scholars should advocate for district-wide adoption of the one approach to achievement for low-income children that evaluators have found to be effective: small class size in the lower primary grades (K-3). Summer Scholars should support all fund-raising initiatives that further this objective.

] Background and Rationale for the Project]

The 1999-2000 school year marked the third and final year of the Stedman Literacy Pilot, which exposes inner-city elementary students to nine months of enhanced literacy instruction and measures whether they demonstrate an accelerated pace of achievement. Initiated in September 1997, the project was conducted collaboratively by the teachers and staff of Stedman Elementary School and staff and volunteers affiliated with Summer Scholars, an independent, non-profit, community-based organization. Summer Scholars raised the money for the Stedman Literacy Pilot and recruited classroom volunteers. Stedman reallocated its budget and hired additional teachers and paraprofessionals to reduce the student/staff ratio. The \$100,000 per year project was funded by the City of Denver's Great Kids Initiative, the Temple Hoyne Buell Foundation, Piton Foundation, Donnell-Kay Foundation, and a variety of private donors.

The Stedman Literacy Pilot incorporates several elements of the successful six-week summer literacy program, Summer Scholars, in a regular school year setting. Its core feature is the conduct of daily literacy classes with first, second, and third graders who were grouped according to reading ability and staffed with teachers, reading assistants, and volunteers to achieve a low student-staff ratio. During its three years of operation, the Pilot included the following:

-] The creation of a daily two-hour uninterrupted literacy block for all first and second graders in 1997-1998 and a 2.5 hour literacy block for first, second, and third graders in 1998-1999 and 1999-2000.

-] The retention of four part-time literacy teachers and paraprofessionals and up to a dozen classroom volunteers to reduce the student/teacher ratio during the morning literacy block to 18:1 and, when fully staffed with paraprofessionals and volunteers, the student/staff ratio to 6:1.
-] The assignment of all first, second, and third grade students to relatively homogeneous groups according to reading ability during the daily literacy block.
-] Reductions in pull-outs for ESL classes and other rival activities during morning literacy classes to insure daily exposure of all students to the literacy block.
-] The provision of staff training on state-of the art teaching techniques through the retention of a part-time lead literacy specialist, who conducted regular classes on balanced literacy and other specific aspects of literacy instruction.
-] The adoption of an effective instructional strategy known as the balanced literacy approach to reading and writing. Components of the balanced approach used at Stedman include reading aloud, shared reading, guided reading, modeled/collaborative writing, phonics, and independent reading and writing.
-] The organization of two family literacy events at the school to involve parents more fully in their children's education.
-] The conduct of an extracurricular reading project to encourage home reading ("Million Minutes") whereby students earn prizes for reading.

Reading tests administered to Stedman students at the start and end of the first year of the Literacy Pilot Project (1997-1998) showed that first and second graders made gains in reading skill that exceeded their counterparts in most other DPS schools. They also

exceeded the reading attainment levels for second graders established in a study of over 1,600 students in seven school districts in Colorado. Stedman students achieved similarly impressive achievement patterns during the pilot's second year, 1998-1999, with 90 percent of second graders and 95 percent of third graders reading at or above grade level on the QRI, an individually-administered reading test used in the Multi-district assessment of normed reading levels for second and third graders. Like Summer Scholars teachers, Stedman teachers attributed the exceptional gains Stedman students made to their daily participation in intensive literacy classes characterized by individual and small group reading activities with similarly situated children in low student/staff ratio settings.

The following presents the results of an evaluation of the final year of the Stedman Literacy Pilot Project. Like the previous evaluations of the program, it examines the efficacy of regrouping students and reducing class size for a portion of the school day and gauges the impact of intensive literacy instruction on student achievement over the 36 week school year. This evaluation also examines the cumulative impact on the achievement of students who experienced more than one year of intensive literacy instruction. The evaluation does not address the impact of home visiting and parent involvement due to a lack of systematic record keeping on these components of the program.

] Evaluation Approach]

Classroom Observations

An experienced educational consultant observed each of the 11 literacy classes in late fall of 1999. The researcher focused on these areas: use of balanced literacy components and extended literacy time, writing instruction, the literacy environment, level of intensity, challenge of content, student engagement, and the use of assisting adults. A

summary was submitted to the program directors based on these observations. Additional information was gathered about behaviors of students, teachers, and assisting adults that affect the intensity and quality of the learning experience.

Teacher Focus Groups and Interviews

In December 1999, evaluators conducted a focus group with Stedman literacy teachers, providing an opportunity for them to reflect on their work. Among the topics discussed were: changes to the program in its third year, teachers' instructional practice, the balanced literacy approach, including homogeneous groupings, impact on student learning, and strategies for maintaining the structure of the literacy block after the completion of the pilot. In the spring, evaluators interviewed individual teachers about their experiences in the final year of the literacy program and their plans for the 2000-2001 school year.

Individually-Administered Reading Tests

First, second, and third grade students were given reading tests administered on a one-on-one basis at the start and end of the school year and their performance on these two tests was compared. In these tests, teachers and independent testers assessed the actual reading behavior of students using standard texts that were progressively more difficult. Young children who were not yet able to read unfamiliar material independently were scored on their ability to distinguish a word from a sentence, point in the correct direction while reading, and locate a specific word. Students who were able to read unfamiliar text independently were rated on their oral reading as they read progressively more difficult passages that have been given grade equivalents. A student's score was based on the number of mistakes s/he made, his or her ability to self-correct, and comprehension of the text.

In September 1999, all first graders and several second graders were tested using the Developmental Reading Assessment (DRA). In May 2000, first graders were retested on the DRA, and second graders on the Qualitative Reading Inventory (QRI). Third graders were tested in both the fall and spring using the QRI. The QRI has been used in a multi-district assessment of the reading performance of second and third grade students involving 7 school districts, 71 schools and approximately 1,600 students at each grade level.

Student Scores on Standardized Reading Tests

Like all DPS students, Stedman second and third graders were given the Iowa Test of Basic Skills (ITBS) in October 1999 and in April 2000. We compared their reading scores on the fall and spring administrations of this exam. This conformed roughly to the start and end of the Stedman Literacy Pilot Project and afforded yet another way to measure student progress and performance using standardized, norm-referenced tests. We also reviewed the scores that third graders achieved on the Colorado Student Assessment Program (CSAP), a state-wide literacy test administered in March 2000.

Student Demographics

We reviewed pupil enrollment cards and elicited a limited amount of demographic information on Stedman's first, second, and third graders. This included their race/ethnicity, age, enrollment in special versus regular education programs, household composition, and languages spoken in the home. We were unable to obtain data on individual students' eligibility for free or reduced lunches. As a result, we do not explore the relationship between income and reading performance.

Student Attendance Patterns

We recorded student absences and tardiness. We also kept track of student attrition and mobility, including withdrawals from Stedman and late enrollments.

Comparisons to Students in Other DPS and Non-DPS Settings

To assess how Stedman students performed relative to their peers in and out of DPS, Stedman first, second, and third graders were compared with their counterparts in DPS in regard to their performance on the ITBS and CSAP assessments. In addition, second graders were also compared with 1,622 second graders studied in 71 elementary schools in Denver, Adams, Douglas, Cherry Creek, Jefferson, and Weld Counties as part of the 1997/98 Multi-District QRI II Second Grade Study. We compared Stedman's third graders with 1,281 students studied in the Multi-District QRI II Third Grade Study conducted in 1997-1998.

Comparison to Other Intensive Literacy Programs in the United States

The Stedman Literacy Project represents one model of intensive literacy instruction. Schools serving similar populations of students have adopted other models of literacy instruction in efforts to raise student achievement in reading and writing. This has included programs offering reduced class sizes to primary level students in Tennessee and Wisconsin and intensive reading interventions known as Success for All and Reading Recovery. All of these efforts have been extensively evaluated and those study findings provide other bases of comparison for the results of the Stedman Literacy Pilot.

Copies of the data collection forms used to conduct the evaluation of the Stedman Literacy Pilot appear in Appendix A to this report.

] Student Characteristics]

The 1999-2000 Stedman Project served a total of 188 children. These students were fairly evenly divided among first (34%), second (34%), and third graders (32%). The average student/teacher ratio across the seven classes of first, second, and third graders was 27:1. During the literacy block, when students attended one of eleven different classes, the student/teacher ratio dropped to 17:1. When the Project was fully staffed with reading assistants and volunteers, the student/adult ratio dropped to 6:1.

There were approximately equal percentages of boys and girls served at each grade level. The majority of the students in each grade were African American, for an overall percentage of 79 percent. Another 18 percent were Latino. Two percent of the students were White and one percent described themselves as "other."

English was the only language spoken in almost all of the children's homes (88%); Spanish was the only language spoken in only one percent of the homes. Not surprisingly then, relatively few children were subject to an English as a Second Language (ESL) pullout from class (6%). Based on enrollment data, many of the children lived in single mother households (42%). About a quarter (29%) lived with both parents, and another 10 percent lived with a parent and step-parent. Nine percent of the children lived with a relative other than a parent, and three percent had living arrangements that were described as "other."

Only five percent of the students were in Special Education Classes. Overall, 13 percent of the students had participated in Summer Scholars, ranging from 22 percent of the third graders to 5 percent of the first graders.

Participation in the Free Lunch program is a good indicator of the family's poverty status. Although information for individual students was not provided, aggregate statistics were. Seventy-one percent of the students in first through third grades were eligible for the free lunch program. Percentages ranged from 67 percent in first grade to 74 percent in third. The school-wide percentage of students eligible for free or reduced-price lunches in 1998-1999 was 81.3 percent.

Among first graders, approximately 88 percent had attended kindergarten. About half of those attending kindergarten went to Stedman, another 27 percent attended another DPS school, and the remainder (15%) attended an out-of-district, out-of-state, or private kindergarten.

Finally, 47 percent of the second graders and 66 percent of the third graders were part of the Stedman Intensive Literacy Program during the 1998/99 school year. Just over half (56%) of the third graders were part of the program during both first and second grades. By May 2000, these 34 students had experienced three consecutive years of heightened literacy instruction. Their performance is discussed in the section of this report entitled "Cumulative Achievement Results."

Table One
Profile of Stedman First, Second, and Third Grade
Students

	First Graders (N=64)	Second Graders (N=63)	Third Graders (N=61)	Totals (N=188)
Gender				
Boys	55%	60%	48%	54%
Girls	45%	40%	52%	46%
Race/Ethnicity				
African-American	77%	80%	81%	79%
Latino/Hispanic	19%	18%	17%	18%
Anglo	4%	2%	2%	2%
Other	0%	0%	0%	1%
Languages spoken in home				
English	85%	91%	87%	88%
English and Spanish	12%	9%	13%	11%
Spanish only	3%	0%	0%	1%
Subject to ESL Pull-Out	3%	7%	9%	6%
Who child lives with				
Mother and father	23%	33%	28%	29%
Mother only	51%	40%	36%	42%
Mother and step-father	7%	7%	13%	9%
Father only	5%	7%	11%	8%
Father and step-mother	0%	2%	2%	1%
Relative	11%	6%	9%	9%
Other	3%	5%	1%	3%
Eligible for Free Lunch	67%	71%	74%	71%
Percent in Special Ed	5%	5%	3%	5%
Ever in Summer Scholars Summer Program	5%	11%	22%	13%

] Student Attendance Patterns]

Student mobility and absenteeism are tremendous problems in urban school systems, and Stedman is no exception. Overall, 17 percent of the students did not participate in a full school year of intensive literacy at Stedman. Approximately 10 percent enrolled late (after September), and 7 percent left before the end of the year (before May). As Table 2 indicates, those who attended only a partial year were typically enrolled for about 149 days, and half were enrolled for less than 145 days.

Table Two
Attendance Pattern for Students Enrolled at Stedman
for Less Than a Full Year

	First Graders (n=7)	Second Graders (n=13)	Third Graders (n=11)	Total (n=31)
Percent enrolling late (after September)	9%	14%	10%	11%
Percent leaving early (before May)	3%	8%	8%	6%
Either late enrollment or left early	12%	22%	18%	17%
Mean (average) days attended	146	155	144	149
Median days attended	141	145	144	145

Students who were enrolled for the full year were absent, on average, 8.9 days. Half of the students were absent for seven or fewer days. Overall attendance was poorest among third graders. They missed an average of ten days, while second graders missed an average of 7.9 and first graders 8.9. The proportion of students who are frequently absent increases with each grade level; 7% of first graders, 10% of second graders and 14% of third graders missed 20 or more days of school. However, the third grade also had

the highest percentage of students with no absences, at ten percent. Only six percent of the second graders came to school every day, and none of the first graders did. As grade levels increase, so does the variation in student attendance. There were no differences between boys and girls with respect to absenteeism. Nor were there differences among students from one and two parent families with respect to absenteeism.

The rate of student absenteeism was higher during the 1999-2000 school year than it was in 1998-1999, when students missed an average of 7.6 days and half missed six or fewer days. Teachers attribute this change to a lack of clarity regarding which staff member was responsible for contacting the parents of absent students. In 1998-1999 this responsibility fell to the family resource coordinator. In 1999-2000 Stedman did not have a family resource coordinator, and the procedure for contacting the parents of children who were frequently absent was followed less consistently.

Table Three
Absences Among Students Enrolled at Stedman
for the Full Year

	First Graders (n=56)	Second Graders (n=49)	Third Graders (n=49)	Total (n=154)
Mean (average) days absent	9.8	8.1	10.0	9.3
Median days absent	7.5	6.0	8.0	7.0
Range of days absent	1-63	0-24	0-44	0-63
Percent with no absences	0%	6%	10%	5%
Percent absent 20 or more days	7%	10%	14%	10%

Second graders were tardy the most often, and third graders the least often. First graders were late, on average, 6.2 days. For second and third graders, the comparable figures were 9.8 and 4.3. The percent who were never tardy increases from first to third

grade. Among Stedman first, second and third graders, the percent who were punctual rose from 18 to 24 to 41, respectively. There is no clear trend in tardiness over the last two years of the Project. While the percent of punctual students increased in the third year of the program, the percent of habitually tardy students worsened.

Table Four
Tardiness Among Students Enrolled at Stedman for the Full Year

	First Graders (n=56)	Second Graders (n=49)	Third Graders (n=49)	Total (n=154)
Mean (average) days tardy	6.3	7.8	4.3	6.1
Median days tardy	3.5	5.0	1.0	2.0
Range of days tardy	0-42	0-51	0-46	0-51
Percent never tardy	18%	22%	41%	27%
Percent tardy 20 or more days	5%	12%	6%	7%

] The Literacy Block]

Reading Instruction

As they have in past years, teachers followed the structure of the balanced literacy model and at the same time gave their own creative spins to the work. The key component of the literacy block is guided reading, and the objective is to expose each child to a daily 25 to 30 minute reading session in a small group of similarly situated students. Unfortunately, the high turnover and inconsistent attendance among reading assistants and volunteers during the 1999-2000 school year forced teachers to incorporate more independent or large group work into the literacy block. These are accommodations which served to lessen the intensity of instruction. In addition to guided

reading, literacy blocks included share reading, independent reading, and “read alouds.” Students rotated to phonics/ spelling stations, independent writing, and occasionally a listening station with books on tape. Teachers also incorporated shared writing into some literacy blocks.

Writing Instruction

Last year the Stedman teachers, reading assistants, and volunteers identified writing as a priority for ongoing professional development. This year the teachers consistently integrated more and better writing activities into the literacy block.

Ingredients for a successful writing program include having challenging and engaging writing activities and well-trained and responsive adult support at all stages of the writing process. In many of the Stedman classrooms, the first condition was met. However, the high turnover among assistants and difficulty in recruiting volunteers once again left many teachers short staffed. With the teachers occupied with guided reading, students often completed work independently at writing centers, and many lacked the support they needed to complete the tasks successfully. While teachers have increasingly provided visual resources to support student writing, such as word walls, many students did not use them consistently during independent work. Teachers also reported that the lack of adult support in their classrooms forced them to offer fewer editing, publishing, and conferencing activities.

] Student Achievement]

During the first month of the 1999-2000 school year, all of the first grade students and several of the lowest achieving second graders at Stedman were tested using the DRA. The majority of the second graders and all of the third graders were tested using the Qualitative Reading Inventory (QRI). Both tests are individualized reading assessments and both have been widely used with Title I and other DPS students. The QRI has been the subject of a multi-district project aimed at developing norms for second and third grade reading. As Table 5 indicates, almost all (86%) of the first graders were below grade level at the start of the year according to the DRA. In 1998-1999, this was the case for 91 percent of first graders. Second grade students were better off. Initial assessments showed only 39 percent of the second graders below grade level when the year began. In 1998-1999, an identical 40 percent entered second grade testing below grade level. Among third graders, 16 percent started the year below grade level according to the QRI. In 1998-1999, 47 percent of third graders started off testing below grade level. These patterns show wider variation in initial reading ability than has been shown in previous years, with fewer children beginning the year reading below grade level with each successive grade, and big improvements for third graders, some of whom experienced two years of literacy block instruction.

When the DRA or QRI was readministered at the end of the school year, students in every grade level showed significant improvements. By the end of the year, only 33 percent of first graders, 14 percent of second graders, and two percent of the third graders were reading below grade level.

Table Five
Reading Level at End of School Year Compared to
Start of School Year on DRA or QRI^W

	First Graders]	Second Graders]	Third Graders]	Total]
Below grade level in Fall 1999	86%	39%	16%	47%
Below grade level in Spring 2000	33%	14%	2%	16%
Number of students	(49)	(51)	(49)	(149)

W Includes only students tested at both the start and end of the school year.

] Differences between start and end are statistically significant.

Did students who began the year reading below grade level move up? The answer appears to be “yes.” Table 6 shows that only 34 percent of the students who started the year below grade level were still below grade level at the end of the year. Most students who began the school year reading below grade level made improvement and ended the year reading at (36%) or above (30%) grade level passages. All of the students who began the year at or above grade level remained at or above grade level in the Spring. The analysis is based on 149 children tested at both the start and end of the school year.

Table Six
Comparison of Student Reading Status at Start and End
of School Year^W

Reading Status of Students at End of Year	Reading Status of Students at Start of Year		
	Below Grade Level	At Grade Level	Above Grade Level
Below grade level	34%		
At grade level	36%	28%	6%
Above grade level	30%	72%	94%
Number of students	(70)	(30)	(50)

W Includes only students tested at both the start and end of the school year.

The QRI indicates the passage level at which students are “instructional” in reading and are handling material that is neither too easy nor too difficult. As Table 7 indicates, second and third graders showed similar patterns of growth, with most students gaining one or two levels.

Table Seven
Number of Reading Passage Levels Gained
by Second and Third Grade Students^W

Number of QRI Passage Levels Gained	Second Graders	Third Graders	Total
None	9%	6%	7%
One	28%	45%	37%
Two	50%	33%	41%
Three	11%	14%	13%
Four or more	2%	2%	2%
Number of students	(46)	(49)	(95)

^W Includes only students tested at both the start and end of the school year.

At what grade level did second and third graders end the year? Table 8 shows that for those students tested on the QRI at the start and close of the year, there were significant gains in passage level. Second graders generally began the year reading at a first grade level and ended the year reading at the level expected of third graders. Third graders began the school year reading somewhat below the level expected of third graders, but ended slightly above the level for fourth graders.

Table Eight
Passage Level at Which Second and Third Grade
Students Read
at End of Year^W

Passage Level at End of Year	Second Graders	Third Graders
First grade	8%	0%
Second grade	20%	2%
Third Grade	43%	22%
Fourth grade	22%	33%
Fifth or sixth grade	6%	43%
Average passage level at start of year	1.3	2.7
Average passage level at end of year]	3.0	4.3
Number of students	(46)	(49)

W Includes only students tested at both the start and end of the school year.

] Differences between start and end of year are statistically significant for second and third graders at .00 level

In addition to the DRA or QRI, there are two other measures of student achievement for which we have information, the ITBS and CSAP.

The Iowa Test of Basic Skills (ITBS) is a standardized reading test administered to second and third graders at the beginning and end of each school year. Table 9 shows that 63 percent of second graders and 85 percent of all third graders scored below grade level at the start of the 1999-2000 school year.¹ On average, at the start of the year, second grade scores on the ITBS showed them to be reading at a level of first grade, seventh month (1.7). Third graders started the year reading at a second grade, third month level (2.3), according to the ITBS. To be at grade level on the ITBS, students in the Fall would have to score at 2.1 and 3.1 respectively.

¹First graders were tested in the Spring of 2000 only, and scored in the 52nd percentile nationally.

By the close of the year, the average second grade Stedman student read approximately at grade level on the ITBS: grade two, seventh month (8.7). The number of second graders with ITBS scores below grade level dropped from 63 percent to 57 percent,

Third grade students began the year reading at grade two, third month (2.3) and ended at grade three (3.0), but a similar proportion (81% to 85%) continued to score below grade level on the ITBS. This means that while reading improved somewhat, most third graders were reading almost a year behind grade level by the close of the year. To be considered at grade level, second and third graders should be reading at 2.8 and 3.8, respectively, at the ITBS administered in the Spring.

Table Nine
ITBS Scores Fall 1999 and Spring 2000^W

	Second Graders	Third Graders
Percent below grade level in Fall 1999	63%	85%
Average grade equivalency score in Fall 1999	1.7	2.3
Number of students	(38)	(40)
Percent below grade level in Spring 2000	57%	81%
Average grade equivalency score in Spring 2000	2.7	3.0
Number of students	(51)	(54)

^W Includes all tested students. To be at grade level in the Fall, second and third graders must score at 2.1 and 3.1, respectively. In the Spring, they must score at 2.8 and 3.8.

ITBS scores traditionally drop for third grade students. The third grade is the first year that the test is timed and students read the questions themselves; in lower grades the teacher reads the questions aloud to the class. Third graders across the district score lower than students at every other grade level, with the average third grader in DPS scoring in the 35th percentile, as compared with 47th and 44th percentile scores for second

and fourth graders, respectively. Third grade performance may also be affected by teacher turnover and quality factors. Although four of the five Stedman teachers who taught grades 1 and 2 had been with the Project for all three years, the third grade teachers were less experienced at teaching third grade and/or new to the Project.

The CSAP, Colorado Student Assessment Program, was introduced in 1998/99 for third and fourth grade students in reading only. The test is standardized and requires students to read text, answer questions, and write sentences. As Figure 1 indicates, among Stedman third graders who were tested, only 13 percent scored “unsatisfactory” on the CSAP test. Another 52 percent scored as “proficient” and 35 percent scored “partially proficient.”

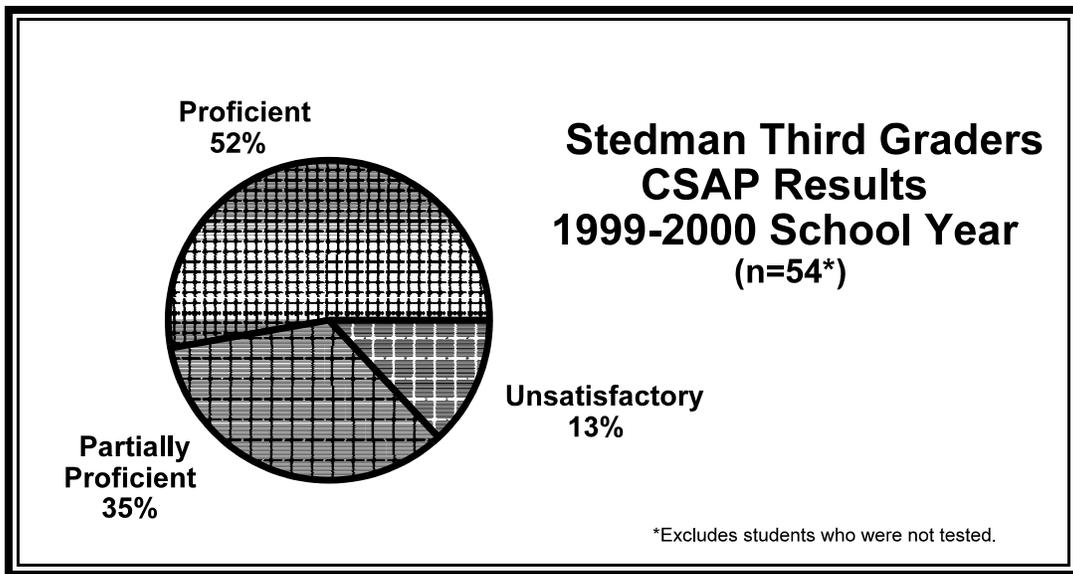


Figure One

Table 10 summarizes the performance of Stedman students on the various tests for which we have information: individually-administered assessments like the QRI (DRA for first graders), the norm-referenced ITBS, and the CSAP. It clearly shows that Stedman students score most favorably on individually-administered tests, with successively higher percentages of students at each grade successfully reading material considered to be

grade level. Their performance on norm-referenced tests, however, raises questions about their grade-level status since the percentage scoring at or above grade level on the ITBS is considerably lower, especially among third graders. Although third graders do better on the CSAP than they do on the ITBS, their performance on that exam also falls short of their extremely favorable performance on the QRI.

These patterns highlight the difficulty in accurately assessing reading skill. While each test has its strengths and weaknesses, none has been correlated with each other and none constitutes a reliable measure of actual grade level performance.

Table Ten
Comparison of Year-End Results for Stedman Students
on Various Reading Tests^W

Percent of Students Testing Below/At/Above Grade Level	QRI/DRA	ITBS	CSAP
First Graders			
Below	33%	69%	
At/Above	67%	31%	
Number of students	(51)	(52)	
Second Graders			
Below	21%	57%	
At/Above	79%	43%	
Number of students	(57)	(51)	
Third Graders			
Below	6%	82%	48%
At/Above	94%	18%	52%
Number of students	(55)	(54)	(54)

^W Includes all tested students, even if not tested at both Fall and Spring.

] Factors Predicting Student Achievement]

There is some evidence that attendance influences which students make progress. Students who began the year reading below grade on the DRA/QRI fared somewhat better on the post-test if their attendance was good. A quarter of the students who began the year below grade level missed 0-4 days, while another quarter missed 12 or more days. Table 11 shows the post-test results for these two groups.

Table Eleven
Gains Made by Students in Grades 1 Through 3
Who Started the Year Below Grade Level by Attendance

	Absent 0-4 days	Absent 12+ days
DRA/QRI rating at the end of the year		
Below grade level	28%	38%
Moves to at or above grade level	72%	62%
Number of students	(18)	(16)

Another correlate of improvement is extracurricular reading. Students who began the year reading below grade level benefitted from time spent on extra-curricular reading. Below grade level students who read at least 500 minutes during the year were significantly more likely to wind up reading at or above grade level (see Table 12). This underscores the importance of having classroom teachers promote extracurricular reading. On average, second graders recorded 640 minutes of extracurricular reading, while first and third graders recorded 390 and 368 minutes, respectively. As part of the

Million Minutes Project, students received incentives for reading at home. Those who read 1,000 minutes or more were more likely to achieve favorable year-end test results than their non-reading counterparts. The 1998-1999 evaluation even showed that readers were dramatically more likely to score at or above grade level than non-readers on the ITBS, QRI, and CSAP. Unfortunately, this year there was no information on student participation in Million Minutes for many students, so it was impossible to assess the impact of extracurricular reading on performance on the ITBS and CSAP.

Table Twelve
Relationship Between Extra-Curricular Reading and QRI
Test Results
for Students Starting the Year Reading Below Grade
Level

Percent of Students Ending the Year Reading . . .	No Extra-Curricular Reading	500+ Minutes Extra-Curricular Reading
Below grade level]	52%	25%
At/Above grade level	48%	75%
Number of students	(23)	(24)
	No Extra-Curricular Reading	1,000+ Minutes Extra-Curricular Reading
Below grade level]	52%	23%
At/Above grade level	48%	77%
Number of students	(23)	(13)

] Differences between the two reading groups are significant at .1.

Not surprisingly, the strongest predictor of achievement is student reading proficiency at the beginning of the school year. For example, for third graders, scoring at grade level on the ITBS at the end of the year was strongly associated with *starting* the year at grade level. Among third graders reading at or above grade 3.1 at the start of the year (e.g., grade level), only 20 percent were below grade level on the Spring ITBS. For those

reading below grade 3.1 on the Fall ITBS, almost 91 percent scored below grade level the following Spring. Unfortunately, while most third graders ended the second grade reading at grade level (e.g., 2.8) according to the Spring 1999 ITBS, they experienced some loss of reading skill over the summer. As a result, they began the third grade with an ITBS of grade two, third month (2.3). Many studies find that low-income students experience a dramatic loss in reading skills over the summer vacation (Cooper, et al., 1996).

] The Results in Context]

The Stedman Literacy Pilot was conducted in a single DPS elementary school; it involved all first, second and third graders. Children were not assigned to the Pilot on a random basis and there is no control group against which their performance may be compared. Despite these limitations, however, there are a number of ways in which the performance patterns evidenced by Stedman students may be contrasted with other relevant first, second, and third graders both within and outside of Denver Public Schools.

Multi-District QRI Study

As in past years, we use comparative data from a Colorado Multi-district Study conducted during the 1996/97 and 1997/98 school years (Felknor and Winterscheidt, 1998; Felknor, et al., n.d.). Aimed at developing regional norms on reading attainment among second and third graders, the study involved administering the QRI to over 1,600 second graders and 1,383 third graders drawn from 71 different elementary schools in seven Colorado school districts. Data collection occurred two to five weeks after the beginning of school and again 2.5 to 5.5 weeks before the end of school, a time period that was identical to the testing period at Stedman. The primary purpose of both testing projects was to

determine the highest passage that each student could read at the instructional level, indicating that the passage was neither too easy nor too difficult. The analysis gauged student improvements in reading skill during the school year as measured by gains in reading passage levels. It also assessed the proportion of second and third graders who meet and exceed standards set by the norming group as well as the proportions that are somewhat below and those that have major reading deficiencies.

Tables 13 and 14 compare Stedman second and third graders with students from the Multi-district study. The results show that, relative to the larger study population, Stedman second graders were reading slightly higher passage levels. For example, at the beginning of the year, 56 percent of the Multi-district Study students were reading at the pre-primer level or below, compared to only 33 percent of the Stedman second graders. This is perhaps due to the fact that many Stedman second graders had participated in the Literacy Pilot Project the previous year as first graders and had benefitted from an enhanced instructional experience. By the end of the year, Stedman second graders continued to show marked differences from the Multi-district Study students. While a quarter of the students in the multi-district study remained at or below the primer level at the end of the year, none of the Stedman students did. Further, at the end of the second grade, more than two-thirds (69%) of Stedman second graders were reading at or above a level 3, while less than half of the students in the Multi-district Study were reading at that level.

Table Thirteen
Stedman Second Graders Compared to
QRI Multi-District Study Second Graders^W

Passage Level	Beginning of School Year		End of School Year	
	Multi-District Study	Stedman	Multi-District Study	Stedman
Primer or below	56%	33%	25%	0%
Level 1	20%	26%	23%	11%
Level 2	6%	18%	7%	19%
Level 3	10%	18%	19%	40%
Level 4+	7%	4%	26%	29%
Number of students	(1,527)	(54)	(1,622)	(53)

W Includes all tested students

Among third graders a similar pattern emerges. At the start of the third grade, Stedman students were reading at slightly higher levels relative to the Multi-district Study students (reflecting the fact that many had participated in the Literacy Pilot Project in the previous year as second graders). For example, 25 percent of the Multi-District Study students, compared to only nine percent of the Stedman third graders, were reading at the primer level or below. By the end of the third grade, 47 percent of the Multi-district Study students, and 75 percent of the Stedman third graders, were reading at level four or higher. Clearly, in both the second and third grades, Stedman students showed greater gains on the QRI than their counterparts in the Multi-district Study.

Table Fourteen
Stedman Third Graders Compared to
QRI Multi-District Study Third Graders^W

Passage Level	Beginning of School Year		End of School Year	
	Multi-District Study	Stedman	Multi-District Study	Stedman
Primer or below	25%	9%	7%	4%
Level 1	23%	9%	7%	0%
Level 2	7%	23%	8%	2%
Level 3	19%	38%	31%	20%
Level 4+	26%	21%	47%	75%
Number of students	(1,622)	(56)	(1,383)	(56)

W Includes all tested students.

DPS ITBS and CSAP Patterns

Another way of assessing the success of the Literacy Pilot Project is to compare Stedman students with their counterparts in other Denver Public Schools on the tests for which comparable information is available. Table 15 presents ITBS scores for three groups of students in Grades 1 through 5 before and after the start of the Pilot in 1998. The groups are Stedman students, all tested DPS students, and tested Black students in DPS. In October 1999, 81.4 percent of Stedman students were African-American and 97.5 percent were classified as non-White.

The comparison shows that Stedman's first and second graders have steadily improved on the ITBS scoring at the 42nd and 45th percentile, respectively, in 2000, which was close to the national average. Stedman's second graders have performed much more favorably on the ITBS in the years since the Literacy Program began than they did previously. The average second grader scored at the 30th percentile in the spring of 1997

(before the program started). Following the introduction of the program, second graders moved up to the 37th, 54th, and 45th percentiles in 1998, 1999, and 2000, respectively.

Like their counterparts in DPS as a whole, students in grades 3 through 5 scored less favorably on the ITBS, suggesting that they lose ground as they move into higher grades. In the third grade in Spring 2000, they scored at the 24th percentile, and in the fourth and fifth grades, they scored at the 36th and 34th percentile, respectively. This tracked closely with ITBS performance patterns for Black students in DPS as a whole.

**Table Fifteen
National Percentile Reading Scores on the ITBS for DPS
Students,
Black DPS Students, and Stedman Students, 1997 - 2000**

	Spring 1997	Spring 1998	Spring 1999	Spring 2000
Grade 1				
All DPS	50	50	52	52
Black DPS	45	41	47	45
Stedman	33	not administered	53	42
Grade 2				
All DPS	44	47	47	47
Black DPS	34	38	38	41
Stedman	30	37	54	45
Grade 3				
All DPS	33	35	36	35
Black DPS	25	27	25	26
Stedman	31	28	23	24

Table Fifteen
National Percentile Reading Scores on the ITBS for DPS
Students,
Black DPS Students, and Stedman Students, 1997 - 2000

	Spring 1997	Spring 1998	Spring 1999	Spring 2000
Grade 4				
All DPS	41	43	41	44
Black DPS	31	33	33	35
Stedman	34	33	38	36
Grade 5				
All DPS	41	43	42	43
Black DPS	32	33	35	36
Stedman	24	38	28	34

While Stedman students at grades 3 through 5 perform no better than their DPS counterparts on the ITBS, they do demonstrate more improvement from Fall to Spring at ever grade level. Table 16 shows that the average second grader advanced 7.7 “percentage points”² compared with the district average gain of 3. While third graders only advanced 2.7 points during the school year, their DPS counterparts fell behind by 5.8 points. More strikingly, fourth graders advanced 9.9 percentage points while their DPS counterparts fell slightly (-0.4). Although the literacy pilot did not include fourth graders, some members of this cohort had been exposed to the pilot during grades 1-3. Finally, even Stedman fifth graders registered more gains than the district as a whole and moved up 6.5 points while their district-wide counterparts fell behind by 1.5. This suggests that the Pilot may have influenced the whole climate of instruction at Stedman.

²The analysis is based on subtracting average fall 1999 and spring 2000 NCEs (Normal Curve Equivalents) on the ITBS which is a mathematically valid procedure and is equivalent to computing a percentage difference.

Table Sixteen
Changes in ITBS Reading Scores (NCEs) Fall 1999 and
Spring 2000 at Stedman and for DPS as a Whole

	Stedman			DPS		
	Fall NCE	Spring NCE	Change	Fall NCE	Spring NCE	Change
Grade 2	39.2	46.9	+7.7	42.9	45.9	+3
Grade 3	38.3	41.0	+2.7	40.5	34.7	-5.8
Grade 4	36.3	46.2	+9.9	42.9	42.5	-0.4
Grade 5	38.9	45.4	+6.5	41.7	40.2	-1.5

The snapshot of student performance at Stedman afforded by the CSAP is also encouraging, with both third and fourth grade students in 2000 registering significant improvement over 1999 scores. In Spring 1999, only 37 percent of Stedman’s third graders scored proficient or better on the CSAP, a proportion that fell below the district average of 43 percent. In Spring 2000, however, 48 percent of Stedman third graders earned a proficient rating. This exceeds the district average of 47 percent. It was substantially higher than performance ratings for Black students in DPS (39%) and for adjacent schools that serve populations with similar demographics which had proficiency ratings that ranged from 23-27 percent.

Stedman fourth graders also made substantial improvement on the CSAP from 1999 to 2000, with the percent scoring at or above proficiency rising from 29 to 35. This mirrored improvement patterns for the district as a whole, and was higher than proficiency patterns demonstrated by Black students in DPS.

Table Seventeen
Percentages of Third and Fourth Graders Scoring At or
Above Proficiency on CSAP Reading Assessment, 1997-2000

	Stedman Students	DPS Students	DPS Black Students
Third Graders			
1998	40%	46%	37%
1999	37%	43%	34%
2000	48%	47%	39%
Fourth Graders			
1997	36%	33%	24%
1998	29%	32%	25%
1999	29%	31%	23%
2000	35%	38%	30%

All students were assigned to one of the following categories regardless of whether they were tested: proficient, partially proficient, unsatisfactory, and not tested.

] Cumulative Achievement Results]

As the Stedman Literacy Pilot draws to a close, we have the opportunity to examine the cumulative effect of three years of intensive literacy instruction on student achievement. The high mobility rates of the families in the Stedman neighborhood presents a challenge to this analysis; only 32 of the 61 third graders attended Stedman for first, second, and third grades; 15 attended Stedman for grades 2 and 3; 14 were new to the program in the third grade.

Given the small sample sizes, the following analysis is merely exploratory in nature. The general pattern was for third graders who attended Stedman for both first and second grades to begin and end the school year reading at higher levels than those who did not. Only 10 percent of the students who experienced all three years of the pilot began third grade reading below grade level, compared to 38 percent of those attending second and third, and 40 percent of those new to the program in third grade. By the end of the year, only 3 percent (representing one student) were below grade level, compared to 7 percent of those attending second and third grades at Stedman and 11 percent of those new to Stedman in the third grade (see Table 18).

Table Eighteen
Achievement Patters on QRI of Third Graders
by Previous Years in Intensive Literacy Program ^W

	Attended Grades 1-3	Attended Grades 2-3	Attended Grade 3 Only
At Program Start			
Below grade level	10%	38%	40%
At grade level	23%	8%	30%
Above grade level	67%	54%	30%
Number of students	(30)	(13)	(10)
At Program End			
Below grade level	3%	7%	11%
At grade level	12%	29%	53%
Above grade level	84%	64%	56%
Number of students	(32)	(14)	(9)

^W Includes all tested students.

Comparing QRI reading levels of students with three years of exposure to intensive literacy at Stedman to the results of the Multi-district Study also suggests favorable results from continued exposure. Table 19 indicates that at the end of the third grade,

only 47 percent of the Multi-district students read at QRI passage level 4 or higher, compared to 85 percent of the Stedman students with three years of intensive literacy.

Table Nineteen
Third Grade QRI Passage Levels from the Multi-District Study and Stedman Third Graders with Three Years of Program Exposure

Passage Level	Beginning of Year		End of Year	
	Multi-District Study	3 Years of Stedman	Multi-District Study	3 Years of Stedman
Primer or below	25%	3%	7%	3%
Level 1	23%	3%	7%	0
Level 2	7%	26%	8%	0
Level 3	19%	42%	31%	12%
Level 4+	26%	26%	47%	85%
Number of students	(1,622)	(31)	(1,383)	(32)

While these findings track with other studies showing that multiple years of exposure to lower class size and intensive literacy efforts pay off, we cannot rule out the possibility that other factors like familiarity with the reading passages used in the QRI or greater family stability helps to explain who receives multiple years of exposure and reading achievement. For example, third grade students from two-parent (step or biological) homes are more likely to have attended all three years of heightened literacy instruction at Stedman (65%) than are students in single parent homes (44%). It will clearly take more research with larger samples to sort out the effects of multiple years of exposure and family stability on reading achievement. We were also unable to assess the impact of family income on achievement since we lacked information on student eligibility for free or reduced-price lunch.

Nor is there any evidence that multiple years of exposure to the intensive literacy program result in higher scores on the ITBS or the CSAP. In every exposure category, virtually identical proportions of students scored below, at, and above grade level.

It is clearly more difficult to get students to successfully apply their improved ability to read and comprehend short passages to other material and in less familiar testing formats. By the higher primary grades (3 through 5), when standardized tests become harder and more independent, fragile readers may be even more disadvantaged than their younger counterparts and score less consistently on different types of tests.

Table Twenty
Comparison of ITBS and CSAP Reading Scores for Students
with Different Amounts of Exposure to the Literacy Pilot

	Attended Grades 1-3	Attended Grades 2-3	Attended Grade 3 Only
CSAP (Spring 2000)			
Below grade level	53%	53%	50%
At grade level	35%	33%	38%
Above grade level	13%	13%	12%
Number of students	(31)	(15)	(8)
ITBS (Spring 2000)			
Below grade level	84%	73%	86%
At grade level	6%	13%	—
Above grade level	9%	13%	14%
Number of students	(32)	(15)	(7)

] Summary and Discussion

]

In the third and final year of the Stedman Literacy Pilot, students demonstrated many of the same achievement patterns they evidenced during the first two years. The most notable finding is that students exposed to heightened literacy instruction in smaller classes consistently experience strong and significant gains on reading tests that are administered on an individual basis at the start and end of the school year. More to the point, these gains exceed those demonstrated by students in regular Colorado classrooms using identical reading tests as part of a Multi-district study of over 3,000 second and third grade students in seven school districts. At the end of the school year, two-thirds of first graders, 85 percent of second graders and 98 percent of third graders successfully read passages that are considered to be at or above grade level. Only about half of the students in the Multi-District study demonstrated this level of reading skill.

As in past years, student performance on standardized reading tests was more mixed. While first and second graders performed the best on the ITBS with the average student performing slightly below the national average (42 and 45 percent), Stedman's third graders scored at the top of the lowest quartile (24%) with 81 percent scoring below grade level and the average student demonstrating a lag of one year in reading proficiency. These patterns are consistent with those demonstrated by DPS students as a whole and Black students in the District. One encouraging difference between Stedman students and the District as a whole was Stedman's higher rate of improvement on the ITBS during the 1999/2000 school year. A comparison of scores on fall and spring administrations of the ITBS shows that Stedman students registered strong improvements at every grade while DPS students made more modest improvement at the second grade level and actually fell behind at grades 3-5.

Stedman's performance on the CSAP, Colorado's test to measure student performance relative to a standard, was extremely encouraging with the proportion of third graders testing at or above proficiency rising from 37 to 48 percent. This proficiency level is slightly higher than the district average of 47 percent and substantially higher than the 39 percent proficiency level for the district's Black third-graders.

Our effort to gauge the cumulative effects of the Stedman Literacy Pilot project also yielded mixed results. While those who were exposed to intensive literacy at Stedman for all three years were more apt to read above grade level on individually-administered reading tests as compared with their counterparts who received two and one years of Pilot project instruction, students with multiple years of exposure performed no better on the CSAP or the ITBS. Substantial proportions of third graders in every exposure category received below grade level ratings on the ITBS; nearly half in every exposure category scored partially proficient or unsatisfactory on the CSAP.

We also get a mixed picture when we attempt to discern whether the Literacy Pilot has lasting effects and whether Stedman's fourth graders experience benefits when they transition into regular-sized classrooms. While Stedman's fourth graders scored worse than District fourth graders as a whole on the ITBS and placed in the 36th percentile (but comparable to Black fourth graders), they registered dramatic gains during the school year and improved by nearly 10 "percentage points," as compared with a drop of 0.4 among fourth graders in the District as a whole. And while Stedman's fourth graders improved on the CSAP with 35 percent achieving a proficiency rating as compared with 29 percent in 1999, they still fell below the district average of 38 percent. On the other hand, Stedman's fourth graders out-performed Black students in DPS as a whole, where only 30 percent were at or above proficiency.

No test does a perfect job of measuring what students know or how they perform; it is impossible to reconcile the different pictures of achievement that each test provides.

Our assessment is also weakened by the lack of a matched group of students in regular classes against whose progress Stedman's students may be compared. Still another shortcoming of our evaluation is the small number of students in the Literacy Pilot and its implementation in just one school. All architects of educational reform programs know that implementation varies dramatically in different school settings and schools with high levels of implementation get higher results in achievement. Stedman clearly had some troubles with staff absenteeism and turnover (especially in the third grade) that may well have reduced effects on achievement.

Nevertheless the Stedman findings of strong improvement in reading skill with few parallel gains on standardized tests are consistent with the mixed results reached in evaluations of other achievement programs for low-income students. One of the most relevant is the widely-adopted Success for All (SFA) program, which is an early grade instructional intervention that focuses on intensive literacy instruction. Like the Stedman Literacy Pilot, SFA includes daily reading instruction, homogenous grouping, periodic individual assessment, and family outreach. While different researchers of SFA's impact on reading achievement reach different conclusions, overall, they find that program impact is highest in the early grades, and generally increases with the number of years students participate in the program (Jones , et al., 1997). At the same time, although SFA students improve in reading and outscore their counterparts from comparison schools, grade-level performance remains elusive. For example, an independent evaluation of SFA (Venezky, 1994) found that the biggest effects occur at the kindergarten level and among lowest-achieving first-graders, and while first graders were almost at grade level on standardized reading tests, they fell behind in each subsequent year. Thus, although SFA students consistently scored higher than a matched group of students not exposed to SFA, only 12.5 percent of SFA students read at or near grade level after five years of program participation.

Evaluations of Reading Recovery, another intensive, highly structured reading intervention for “at risk” elementary school students also show strong boosts in reading skill but smaller effects over time. Based on individualized tutoring and phonics instruction, evaluators find that Reading Recovery leads to improved scores for low achieving readers who might not have succeeded otherwise. Furthermore, on standardized tests, primary school students perform on average, better than similar children who were not enrolled in the program. Effect sizes diminish dramatically with time, however, and by the end of third grade those effects are minimal (Shanahan and Barr, 1995). Due to the high costs of program administration and the relatively low savings generated by low retention rates and special education costs, Shanahan concludes the costs of Reading Recovery outweigh the benefits.

An approach to achievement for low-income children that evaluators find to be highly cost-effective stresses small class size in the primary years. In the Tennessee Class Size Study, Project STAR researchers found clear and consistent achievement benefits for kindergarten through grade three students who were taught in classes with a teacher: student ratio of 1:15. Small classes scored highest on standardized achievement tests in all four grade levels (K-3) and in all locations (urban, suburban, rural) with the greatest gains made by minority students in inner-city settings. The presence of classroom aides did not make a difference in student achievement with large classes with and without aides performing identically and less favorably ([Project STAR], n.d.). Follow-up studies with STAR students find that the advantages of small classes are retained after students return to regular-size classes with the most recent research showing benefits all through high school. As a result of being less likely to be retained and to graduate on schedule, students who had been in small classes during kindergarten through third grade may have saved the state of Tennessee \$3,547,302 in education costs (Pate-Bain, et al., 1999). It should be noted, however, that even in this highly effective program, researchers find the biggest average gains in performance on standardized tests (4 percentile points) in

the first year of assignment to small class, irrespective of grade, with smaller increases thereafter (1 percentile point per year)(Krueger, 1997).

Wisconsin has also reduced class sizes to 15 in grades K through 3 in an effort to improve the achievement of poor students. A study conducted in the third year of the Student Achievement Guarantee in Education program (SAGE) finds that SAGE students scored significantly higher on standardized tests than did their counterparts at comparison schools. Indeed, researchers conclude that during the third grade year, the achievement gap for Black and White students narrowed at SAGE schools while it widened at comparison schools ([SAGE], 1999).

These studies, along with the Stedman evaluation underscore the benefits of intensive educational policies for low-income students at the primary level. Even if they do not always lead to grade-level performance on norm-referenced tests, they consistently produce strong and significant gains in reading skill among the most fragile learners and narrow the achievement gap between White students and racial minorities, on newer, standards-based performance tests like the CSAP. This is an impressive achievement, especially for a program that served 188 students with an annual budget of only \$100,000 (a cost of \$532 per student per year). There is research evidence that more extensive school reform efforts can have an even bigger impact. For example, in Tennessee, where small class size has been used the longest and with the largest numbers, students who experienced small classes in K-3 show positive and significant long-term effects through high school.

] Recommendations]

The Stedman Literacy Pilot ended in June 2000; Summer Scholars will cease to play a formal role in fund-raising, program design, implementation, and evaluation. Fortunately, the program has become firmly rooted in the school; the principal and teachers hope to incorporate elements of the Literacy Pilot in future instructional formats on a long-term basis. The following recommendations are designed to help Stedman staff with program continuity and to help guide Summer Scholars in its advocacy on behalf of Denver's disadvantaged readers

] Collaborations between schools and community groups can be extremely effective ways of generating resources, energy, and focus to support literacy interventions. Through the fund-raising efforts of Summer Scholars, Stedman benefitted from an additional \$100,000 per year, which was used to hire part-time teachers and reading assistants to reduce class size for 2.5 hours each day. Summer Scholars also recruited volunteers to enhance literacy instruction further. Finally, the collaboration led to an enhanced planning and evaluation effort which heightened program visibility and accountability. Although student performance on norm-referenced tests is mixed, the results are extremely impressive on individually-administered reading tests and standards-based performance tests. Schools should actively pursue collaborations with community groups that have expertise in achievement programs for low-income students to augment their efforts to promote student literacy.

] It is recommended that Stedman continue to offer small, homogeneous classes for literacy instruction for at least 2.5 hours per day. All relevant internal and external resources should be marshaled for this effort with the goal of achieving literacy classes with a student/teacher ratio no greater than 15:1. In addition to pursuing

relevant federal, state and local grants, Stedman should continue to explore other possible reorganizations such as the use of Title I reading staff, the Assistant Principal, and other non-classroom personnel, to support regrouping for literacy classes with lower student/teacher ratios.

] It is recommended that Stedman invest its limited resources in teachers rather than reading assistants and other paraprofessionals who have displayed an unacceptable level of absenteeism and turnover. Due to high levels of staff turnover and absenteeism, some literacy classes operated without full staffing. This meant that teachers had to rely more on large-group instruction and could not offer all students a daily, opportunity to read appropriate material in a small group with similarly situated students. The Tennessee STAR Project found no achievement benefits to large classes with or without an aide; the only benefits came from exposure to teachers in low ratio settings.

] Students who attend school regularly make more progress than those who don't and many Stedman students continue to enroll late, leave early and attend irregularly. The loss of the school's Family Resource Coordinator has made lines of responsibility unclear for monitoring attendance and intervening with families. Stedman must implement an aggressive attendance program and identify personnel who will quickly flag children who are repeatedly absent or tardy. Since absenteeism is significantly greater among children in single versus two parent homes, this might also be a population to target for intervention by DPS social workers.

] Whether they begin the year reading at or below grade level, students who engage in extracurricular reading do better on year-end tests. It is recommended that teachers actively support the Million Minutes Project which provides incentives for home reading.

-] While students in the lowest grades make the most gains and intervention should appropriately focus on students in grades K-3, there is a definite need for additional remediation at later grades. Before and after school programs should include one-on-one tutoring, and small group support in reading. Older students also need summer school interventions to prevent the dramatic loss in reading skill that typically occurs for low-income students during the summer months.

-] Three years is not a long enough test for a pilot project aimed at demonstrating the effects of an intensive literacy program. Nor is implementation in a single school setting reliable. Complex programs take time to get right. And since there is always a range in implementation depending on school culture and climate, any test of a new program should be done in multiple settings in order to assess the impact of high and weak program implementations.

-] The Stedman Literacy Pilot Project was implemented with a minimal amount of advance planning and attention to school and community “buy-in.” Other literacy initiatives in the Denver area are dedicating a full year to the planning process. Summer Scholars should monitor these newer efforts and assess the appropriate mix of planning and action before undertaking another school year program.

-] Summer Scholars should advocate for district-wide adoption of the one approach to achievement for low-income children that evaluators have found to be effective: small class size in the lower primary grades (K-3). Summer Scholars should support all fund-raising initiatives that further this objective.

] References]

- Cooper, H., B. Nye, K. Charlton, J. Lindsay and S. Greathouse. 1996. "The Effects of Summer Vacation on Achievement Test Scores: A Narrative and Meta-analytic Review." *Review of Educational Research*, 66 (3): 227-268.
- Felknor, Catherine M. and Victoria Winterscheidt, "Classroom Implications From the Multi-District QRI II Study," *Colorado Reading Council Journal*, Spring 1998.
- Felknor, Catherine, et al., "Cooperative Project for Developing Regional Norms and Information for Establishing and Assessing Attainment of Standards Related to Second Grade Reading - 1996-99 School Year." General information about data collection, scoring, and highlights of the analysis. Unpublished manuscript, no date, Denver, Colorado.
- Jones, Elizabeth M., Gottfredson, Gary D., and Gottfredson, Denise C. 1997. "Success for Some: An Evaluation of a Success for All Program," *Evaluation Review*, 21 (6): 643-670.
- Krueger, Alan B. 1997. "Experimental Estimates of Education Production Functions." www.nber.org/papers/W605.
- Pate-Bain, H., B.D. Fulton, Jayne Boyd-Zaharias. 1999. "Effects of Class-Size Reduction in the Early Grades (K-3) on High School Performance: Preliminary Results from Project STAR. www.nea.org/issues/classsize/bain.
- [Project STAR]. N.d. "Final Executive Summary, Kindergarten Through Third Grade (1985-1989)." www.cde.ca.gov/classsize/eval/projstar.
- [SAGE]. 1999. "Executive Summary: 1998-99 Results of the Student Achievement Guarantee in Education Program Evaluation." University of Wisconsin-Milwaukee. www.uwm.edu/Dept/CERAI/documents/execsummaryyear3.
- Shanahan, T. and R. Barr. 1995. "Reading Recovery: An Independent Evaluation of the Effects of Early Instructional Intervention for At-risk Learners." *Reading Research Quarterly*, 30 (4): 958-996.
- Venezky, R.L. 1994. "An Evaluation of Success for All: Final Report to the France and Merrick Foundation." Newark, NJ: University of Delaware, Department of Educational Studies.

] APPENDIX A]