



Evaluation of Massachusetts Reading First

## Year 5 Evaluator's Report Summary and Conclusions

**A Review of Program Activity: July 1, 2006 - June 30, 2007**

---

Presented to the Massachusetts Department of Elementary and  
Secondary Education

April 2008



## Summary and Conclusion

This report provides an overview of activities and student outcome data for Massachusetts' fifth year of funding under the federal Reading First program. Its primary focus is the analysis of student assessment data and examination of changes in student outcomes for schools with three or four full years of classroom implementation as well as results for schools participating in the state-funded John Silber Reading Initiative.

### Program Description

During the 2006-2007 academic year, 42 districts received funding totaling about \$11 million dollars through the Massachusetts Reading First program. In total, 89 schools participated during this period. Those schools employed nearly 2,100 administrators, reading specialists and K-3 classroom teachers. Through the first four years of classroom implementation (fall 2003 through spring 2007) more than 80,000 Massachusetts K-3 students participated in Reading First. Table 1 provides a snapshot of the characteristics of the K-3 students enrolled in Massachusetts Reading First schools on October 1, 2006.

Total enrolled	24,656
Special Education students	14%
English Language Learners	21%
Low Income students	72%
White students	35%
Hispanic/Latino students	39%
Black/African American students	15%

In addition, 36 schools participated in the state-funded John Silber Reading Initiative, which is modeled after Reading First. The Silber program provides funding to schools that have an identified need, but are not eligible for Reading First, primarily because they don't meet the poverty criteria. Silber schools receive professional development (including foundational training) and support to improve K-3 reading instruction. They are included as part of the Reading First regional network and statewide meetings. They are required to administer the DIBELS and GRADE assessments to their students. Table 2 provides a snapshot of the characteristics of K-3 students enrolled in Silber schools on October 1, 2006.

Total enrolled	8,991
Special Education students	14%
English Language Learners	19%
Low Income students	61%
White students	52%
Hispanic/Latino students	30%
Black/African American students	9%

While individual schools and districts have some flexibility in how they implement their Reading First grants, all must incorporate the following basic program requirements:

- Develop and implement an instructional model centered on tiers of curriculum delivery.
- Employ a full-time reading specialist in each participating K-3 school to provide high-level support to classroom teachers and others involved in the teaching of reading.
- Participate in foundational training as well as ongoing professional development and support provided by the Massachusetts Department of Elementary and Secondary Education.
- Administer designated student assessments and use data to inform instruction.

In addition to specific professional development events, the Massachusetts Department of Elementary and Secondary Education employs a cadre of implementation facilitators whose role is to provide ongoing, direct support to staff in Reading First schools. They work primarily with the reading specialist in each school but also frequently have direct contact with building administrators and teachers. The implementation facilitators also lead bimonthly regional meetings, which bring together reading specialists and district coordinators to share experiences and implementation challenges.

Staff from the Department's Office of Reading also conduct monitoring visits to each Reading First school. The objective is to identify areas of strength and weakness as well as actions needed to improve Reading First implementation. After the visit each school receives a letter summarizing findings from the visit and is expected to work with their implementation facilitator to develop an action plan addressing those findings.

## Student Outcome Measures

The program evaluation utilizes results from three student assessments as the basis for measuring student improvement and providing comparisons among groups of students. They are:

- The DIBELS Oral Reading Fluency subtest is a standardized, individually-administered assessment developed at the University of Oregon. Based on performance, students are placed in three categories – at risk, some risk and low risk.
- GRADE is a norm-referenced, group-administered assessment developed and marketed by Pearson. It is a comprehensive test covering the five key components of reading and offers multiple level tests for use across many grade levels. The Massachusetts Department of Education has established four categories of reading achievement based on students' scores – weak, low average, average, or strength. Students scoring in the average or strength category are considered to be performing "at or above grade level."
- The Massachusetts Comprehensive Assessment System (MCAS) reading test is designed to assess the reading skills of all public third graders in the state. The primary focus of the test is reading comprehension. Results of the third grade reading test are reported in terms of four performance levels – warning, needs improvement, proficient, and above proficient. The above proficient category was new for the 2006 test. For the purposes of the Reading First evaluation those students are grouped in the proficient category.

## Findings

- ❖ As defined by the U.S. Department of Education, the main criteria for evaluating the impact of Reading First is whether the program has resulted in an increase in the percentage of students performing "at or above grade-level" and a decrease in the percentage of students with "serious reading difficulties." To address these criteria, Massachusetts relies primarily on results from the DIBELS Oral Reading Fluency and GRADE

assessments. Results from both of these assessments demonstrate that Massachusetts has met these improvement criteria for all grade-levels and participating cohorts.

DIBELS Oral Reading Fluency results show improvement for all grade-levels and cohorts, including increases in the percentage of students in the *low risk* category, decreases in the percentage of students in the *at risk* category, and increased mean scores.

Results on the GRADE assessment also show increases in the percentages of students scoring in the *average/strength* category (stanine 5-9), decreases in the percentages of students scoring in the *weak* category (stanine 1-3), and increases in mean scores at all grade levels on the over time. Eighteen schools (12 Reading First and 6 Silber) stood out as having 80 percent or more of their students performing at benchmark on the spring 2007 GRADE assessment. Since they began program implementation, about 70 percent of Reading First and Silber schools demonstrated increases in the proportion of students in the *average/strength* category and decreases in the proportion of students in the *weak* category on the GRADE assessment. These included about 30 percent of the schools that showed substantial improvement with *average/strength* increases **and** *weak* decreases of at least 10 percentage points.

- ❖ All of the designated subgroups (special education, limited English proficient, low income, African American/black and Hispanic/Latino students) have shown cumulative improvement as measured by performance on the GRADE assessment. Of particular note are those subgroups with levels of improvement which meaningfully exceed the general population (an indication that the performance gap for these students is narrowing). Those subgroups are: RF cohort 1 first and second grade special education students, RF cohort 1 first grade limited English proficient students, and RF cohort 2 third grade limited English proficient students. There are also a few subgroups with levels of improvement that are meaningfully smaller than the general population (an indication that the performance gap for these students is widening). Those subgroups are all from RF Cohort 1. They are: first and second grade African American students and third grade limited English proficient students.
- ❖ Relative performance on the Reading First assessments (DIBELS ORF and GRADE) and MCAS shows that Reading First students are improving, but so far not enough to yield marked improvement on the more challenging MCAS test. Since baseline, third grade MCAS results for the state as whole, Reading First, and Silber all show decreases in the level of *proficiency* and increases in the percentage of students scoring in the *warning* category. Annual changes in *proficiency* from 2006 to 2007 are more hopeful with stable results for RF cohort 2 and small improvements statewide, for RF cohort 1 and JSER cohort 2. However, during the same period, the percentage of students performing at the *warning* level increased statewide and for each of the RF cohorts.

Yet, when judged by the percentage of students meeting or exceeding the *needs improvement* level (a standard much more consistent with “grade-level” performance on nationally-normed assessments), Massachusetts students perform quite well. In 2007, 91 percent of students statewide met or exceeded the *needs improvement* cut score as did 82 percent of students in RF and JSER schools.

Among participating schools, there are wide disparities in MCAS performance. In 2007, 11 schools (nine RF and two Silber) had third grade MCAS *proficiency* rates equal or better than the statewide rate of 59 percent and 10 (five RF and five Silber) had *warning* rates lower than five percent. At the same time, 22 schools (16 RF and six Silber) had *proficiency* rates of 25 percent or less and 20 (18 RF and two Silber) had *warning* rates of 33 percent or more. Since the year prior to implementation, one-quarter of Reading First schools and about 22 percent of Silber schools demonstrated increases in the proportion of students attaining *proficiency* and decreases in the proportion of students in the *warning* category on the MCAS third grade reading test. These included about six percent of Reading First and Silber schools that showed substantial improvement with *proficiency* increases **and** *warning* decreases of at least 10 percentage points.

- ❖ Comparing the performance of proficient and non-proficient students indicates that focusing on the following issues may further improve MCAS proficiency rates: developing faster and more accurate decoding skills;

practicing with longer and more difficult authentic text – including high-quality expository text; building receptive vocabulary; developing strategies to infer meaning from text; and helping students respond to literature, especially in writing.

- ❖ For each of the included cohorts, the 2006-2007 effectiveness index for *average/strength* students shows that 95 percent of those who began the year in the *average/strength* categories ended the year at that level. Furthermore, about half improved their performance by one or more stanine, including about 30 percent who moved from *average* to *strength*. More than half of all Reading First and Silber schools demonstrated instructional effectiveness for these students of at least 95 percent, including 10 schools at 100 percent. The programs by the most commonly used core publishers (Harcourt, Houghton-Mifflin, and Scott Foresman) all appear to provide highly effective instruction to students who began the school year meeting benchmark on the GRADE assessment. At the second and third grade levels, the data do suggest that in some respects schools using Scott Foresman perform better than those using Harcourt.
- ❖ For all of the included cohorts combined, the 2006-2007 effectiveness index for *low average* students shows that about 70 percent of those who began the year in the *low average* category ended the year in the *average/strength* categories. For each of the cohorts, instruction was the most effective at the first grade level, especially with regard to moving students from the *low average* category to the *strength* category. Nineteen schools demonstrated instructional effectiveness of at least 85 percent, including Sheffield Elementary in Gill-Montague, which moved all of its *low average* students into the *average/strength* categories.
- ❖ For all of the included cohorts combined, the 2006-2007 effectiveness index for *weak* students shows that 56 percent of those who began the year in the *weak* category ended the year in the *low average* category or higher. As with *low average* students, instruction for *weak* students was the most effective at the first grade level, especially with regard to moving students from the *weak* category to the *average* category and even more so in moving students from the *weak* category to the *strength* category. Twenty-two schools demonstrated instructional effectiveness for *weak* students of at least 70 percent.
- ❖ Case studies conducted at schools with promising student outcomes suggest the following:
  - School leaders' active involvement in both policy and execution demonstrate belief in Reading First and foster staff buy-in.
  - Principals' background, school history and incremental approaches to change contribute to schools' experience of success with Reading First.
  - Professional development is ongoing and customized to meet the needs of instructional staff.
  - School culture reflects a commitment to valid data and sophisticated analyses of assessment results. Ongoing assessment and increasingly institutionalized data review practices position teachers to identify children's weaknesses and pinpoint their instruction appropriately.
  - Schools integrate the roles of classroom teachers and interventionists.
  - Teachers strike a balance between fidelity to the curriculum and flexibility to address specific needs.
  - Differentiation addresses the needs of more proficient students as well as struggling readers.
  - Student success feeds teachers' professional satisfaction and leads to steadily increasing expectations for student learning.
- ❖ The vast majority of Reading First staff – 98 percent of reading specialists, 91 percent of teachers and principals, and 87 percent of district coordinators – reported that RF had at least moderately improved their overall knowledge about effective reading instruction. Those figures include about 80 percent of principals and reading specialists who reported that their overall knowledge had been “very much” improved. The perceived impact on knowledge was fairly well distributed across the five dimensions of reading. Although Reading First is often criticized for being too focused on phonemic awareness, phonics, and fluency, reading

specialists perceived the greatest impact in their knowledge of vocabulary. Principals perceived the greatest impact in the areas of fluency and vocabulary.

- ❖ Teachers reported that Reading First had improved their instructional practice, and that perceived improvement was fairly consistent across the five dimensions of reading. The areas in which teachers felt the least improvement were selecting effective curricula and instructional materials, and planning and managing differentiated instruction (though these areas still received relatively high mean scores). Principals and district coordinators reported quite similar impacts on their teachers.
- ❖ Across roles, Reading First staff report that the program has had a moderate to strong impact on students' reading skills, particularly in the areas of phonological and phonemic awareness and phonics and word study. However, only about half of reading specialists think that their schools are using both supplemental and intensive intervention programs effectively, which may indicate that many schools are still struggling with the intervention component of Reading First.
- ❖ The literacy block, full-time reading coach, tiered curriculum delivery and the DIBELS assessment were all perceived as critical success factors of the Reading First program. Aspects of the grant that were perceived to be less important were the particular core curriculum, the foundational reading course, and the GRADE assessment. Nearly all of the district coordinators and principals indicated that once grant funds were no longer available, their schools would be very likely to continue the uninterrupted extended literacy block and using their selected core curriculum. In contrast, though highly valued, substantially fewer schools indicated that they were very likely to continue to employ a full-time reading coach position and tiered curriculum delivery.

After five years of funding, the Massachusetts Reading First program has had positive measurable impacts. Increases in fluency continue to mark an important first step in helping students read and comprehend appropriate text for their grade level. Among Reading First schools, there is meaningful improvement in overall reading ability and many of the cumulative gains over the course of the grant are statistically significant. Although reading specialists perceive that their schools could be doing a better job providing intervention, the effectiveness index data show some success in improving performance of students who begin the year at moderate or substantial risk for reading difficulties. Perhaps most importantly, Reading First staff are generally quite positive about the program's impact on their own knowledge and practice with regard to effective reading instruction. In the long run, that may be the most meaningful impact as it holds the potential to positively impact students' reading skills long after program funding has disappeared. At the same time that it recognizes and celebrates the progress to date, it will be important for the Massachusetts Department of Elementary and Secondary Education to better understand the challenges that limit that improvement, particularly with regard to the MCAS reading test, and provide the necessary professional development and support to move forward.