Background

Reading Recovery—a short-term, one-to-one literacy intervention for first graders—has been used extensively to counter early literacy difficulties. Developed in New Zealand in the 1970s, it has been used in the United States since 1984 and in Massachusetts since 1990. Previous studies have documented the success of Reading Recovery (Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994; Askew & Frasier, 1997; Iversen & Tunmer, 1993; Stahl, Stahl, & McKenna, 1999). Even its most critical reviewers find it to have positive results on students, closing the gap between the weakest readers and their classmates (Shanahan & Barr, 1995; Iversen & Tunmer, 1993). Other studies find lowered retention and special education rates as a result of Reading Recovery, leading to significant cost savings (Wasik & Slavin, 1993; Dyer & Binkney, 1995). But how long term are the effects of Reading Recovery? Do children who successfully complete the intervention retain its benefits?

Researchers have found Reading Recovery instruction has a positive influence on students’ literacy skills through the second grade (Center, Wheldall, Freeman, Outhred, & McNaught, 1995; Askew & Frasier, 1994) and even longer (DeFord, Pinnell, Lyons, & Place, 1990; Wasik & Slavin, 1993; Moore & Wade, 1998; Brown, Denton, Kelly, & Neal, 1999). In this study, the effectiveness of Reading Recovery in Massachusetts was examined using the third-grade Massachusetts Comprehensive Assessment System (MCAS), a state-mandated standardized test and an assessment that has received increased attention with the passage of the No Child Left Behind Act of 2001 (2002). Individual student scores on this test are categorized as Warning, Needs Improvement, or Proficient.

Data

Throughout the 2002–2003 academic year, Reading Recovery teachers collected MCAS data on former Reading Recovery students throughout Massachusetts in order to examine the longitudinal effects of Reading Recovery instruction. These data were submitted to Lesley University for analysis. The sample used for this study included all children from the 1999–2000 Reading Recovery cohort who took the third-grade reading portion of the MCAS exam in the spring of 2002 and for whom teachers submitted data.

The majority of children who successfully complete Reading Recovery lessons continue to perform at grade-level standards.

Of the 1,389 students who participated in Reading Recovery in the 1999–2000 academic year and for whom we have follow-up data, we have retention information on 881 of them. Approximately 9% of these 881 students were retained sometime between kindergarten and third grade. The majority of the children retained do not have MCAS scores as they had not yet taken the third-grade MCAS exam in the spring of 2002. For this reason they have not been included in the study. We have no information on why these children were retained. The remaining 768 students for whom we have retention information progressed normally.
• for whom we had data on their school district, and
• for whom both raw and categorical MCAS score \((n = 938)\) were available. These 938 third graders from 47 school districts represent approximately 20% of all children served by Reading Recovery in Massachusetts in the 1999–2000 academic year.

Although these data were not collected as a random sample but based on teachers’ access to records, the rate of successful completion (67.7%) for children in the sample is very similar to the statewide rate in the 1999–2000 academic year suggesting these data may be representative of the larger Reading Recovery population served in 1999–2000. In this study, the MCAS scores of Reading Recovery students obtained in spring 2002 are compared to data of the 47 school districts from the Massachusetts Department of Education Web site which reports that from these districts a total of 12,840 third graders sat for the MCAS exam.

Limitations of the Data
This study compares the scores of Reading Recovery students to the comparison group students using average scores across school districts. The data were analyzed in this way because although data on former Reading Recovery students were available at the individual level, non-Reading Recovery data were only available in the aggregate or as a group. The lack of comparable individual level data means an inability to perform individual level analysis to compare students within schools or districts. Each district must stand as one case with average scores of former Reading Recovery and non-Reading Recovery students compared across districts.

Further, we do not have data available on important demographic characteristics such as age, gender, race and ethnicity, poverty status, and English language ability. This lack of demographic data means an analysis of sustained effects outcomes for specific subgroups of Reading Recovery students was not possible, nor was a determination of whether such demographic characteristics affected the results reported here.

Moreover, a characteristic of this study is that it reports on the third-grade MCAS scores of approximately 20% of the 1999–2000 Reading Recovery cohort in Massachusetts. This is an important limitation for this study because missing information tends not to be randomly assigned and could thus be producing bias in the results.

Methodology
The need for services and the level of coverage required by districts that implement Reading Recovery varies significantly. Accordingly, the average MCAS scores and distributions among former Reading Recovery students must be statistically weighted by the number of Reading Recovery students served in each district. I have weighted the district level Reading Recovery MCAS averages by the total number of Reading Recovery students in each of the respective districts using the following formula:

\[
X = \frac{\sum \text{(RR district MCAS mean) \times (# of RR students in district)}}{\sum \text{RR students in all districts included}}
\]

I have also weighted the district level MCAS average by the size of the underlying population taking the exam in each district. Using the same formula as above, I have calculated weighted category (Warning, Needs Improvement, Proficient) averages.
and passing rates across the 47 districts for all students, for former Reading Recovery students, and for Reading Recovery students who successfully completed their series of lessons. To determine whether any differences are statistically significant, I have used T-Tests to compare the weighted mean raw MCAS scores, categorical distributions, and passing rates across the 47 school districts to see how all Reading Recovery students as well as successfully completed Reading Recovery students compare with the overall general population in their respective districts.

**Findings: District Level**

Table 1 presents the study's results across the 47 districts for all third-grade students who took the MCAS exam, all Reading Recovery students who received any portion of the intervention, and Reading Recovery students who successfully completed their series of lessons.

The weighted average raw score for all third graders who took the MCAS across the 47 districts is 29.44, as compared with 27.10 for all Reading Recovery students (who may have received as little as one lesson), and 28.36 for successfully completed Reading Recovery students.

Among the outcome categories, 7.29% of all students are in the Warning category, versus 7.14% of all former Reading Recovery students, and 4.51% of successfully completed Reading Recovery students. In short, Reading Recovery students are showing up in the lowest category at lower rates than students overall.

The Needs Improvement and Proficient categories show somewhat different patterns. Approximately 30% of all students versus 45.34% of all former Reading Recovery students, and 40.10% of successfully completed Reading Recovery students are in the Needs Improvement category. At the same time, rates for the Proficient category are 62.78%, 47.34%, and 55.39% respectively for the same groups. The fact that nearly half of all Reading Recovery students and more than half of those who successfully completed the intervention are found Proficient is an important finding given that 2 years earlier they were in the bottom 20% of readers.

Figure 1 shows that scores for Reading Recovery students who successfully completed their series of lessons are more similar in distribution to the general population than are all Reading Recovery students in both the Needs Improvement and Proficient categories. This is expected, given that successfully completed Reading Recovery children were identified as reaching average reading levels, while the sample of all Reading Recovery students includes those who did not receive a full intervention for a variety of reasons or who were recommended for additional services after 20 weeks.

Though Table 1 and Figure 1 give some indication of outcomes by student group, data were further examined to determine whether the differences between groups were statistically significant.

Table 2 on the following page presents the findings from the T-Tests. The mean weighted raw scores of both groups of Reading Recovery students are lower than the overall mean score at a .001 level of statistical significance. The comparison of category outcomes shows a more positive picture, with all Reading Recovery children performing similarly to or even better than the general population in certain instances. Reading Recovery children who successfully completed their
series of lessons show positive, statistically significant differences from the general population. That is, those children are less likely to be in the Warning category. The larger Reading Recovery group, which includes any child who received even one Reading Recovery lesson, does not appear in the Warning category at statistically significantly different rates from the overall population. Given that the Reading Recovery children were originally among the 20% lowest performing readers, this is an impressive finding. Both Reading Recovery groups appear in the Needs Improvement category at statistically significantly higher rates than the overall population. Both groups appear in the Proficient category at lower rates than the general population.

In terms of passing the MCAS exam, Reading Recovery children perform similarly to or even better than the general student population. Children must receive a Needs Improvement or Proficient to pass the exam. The Warning category constitutes failing.

Figure 2 highlights the rate of passing by student group with an overall passing rate of 92.74% as compared with 92.86% and 95.49% for all Reading Recovery children and Reading Recovery children who successfully completed their series of lessons respectively. Reading Recovery children who successfully completed their series of lessons are significantly more likely to pass the MCAS than the overall population (at the .001 level of significance).

**Discussion**

The MCAS exam has become an increasingly popular measurement of a school’s success, with the rate of passing viewed as a barometer of how well school districts are doing. MCAS results are also being used as a method of measuring instructional efficacy. This study was designed to use the MCAS results to measure the efficacy of Reading Recovery instruction in Massachusetts. Based on the results of this study, Reading

---

### Table 2. T-Tests Comparing Student Groups Across 47 Massachusetts School Districts

<table>
<thead>
<tr>
<th>MCAS Outcomes</th>
<th>All Former Reading Recovery Students</th>
<th>Successfully Completed Reading Recovery Students</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N = 938$</td>
<td>$N = 621$</td>
</tr>
<tr>
<td>Mean Raw Score (vs. $x = 29.44$)</td>
<td>27.10*** 2.56 0.08</td>
<td>28.36*** 2.63 0.11</td>
</tr>
<tr>
<td>Percentage Proficient (vs. $x = 62.78$)</td>
<td>47.34*** 18.79 0.61</td>
<td>55.39*** 20.20 0.81</td>
</tr>
<tr>
<td>Percentage Needs Improvement (vs. $x = 29.96$)</td>
<td>45.52*** 16.31 0.53</td>
<td>40.10*** 18.07 0.73</td>
</tr>
<tr>
<td>Percentage Warning (vs. $x = 7.29$)</td>
<td>7.14 6.43 0.21</td>
<td>4.51*** 5.73 0.23</td>
</tr>
<tr>
<td>Percentage Passing (vs. $x = 92.74$)</td>
<td>92.86 6.43 0.21</td>
<td>95.49*** 5.73 0.23</td>
</tr>
</tbody>
</table>

*P < .05, **P < .01, ***P < .001

Statistics on all students are weighted by the number of children in each district who took the MCAS exam. Statistics on Reading Recovery students are weighted by the number of Reading Recovery students served in each district.

Source: Data on school districts’ average MCAS scores are from the Massachusetts Department of Education Web site, where results from specific school districts are available. Data on former Reading Recovery students were collected by Reading Recovery teachers and teacher leaders and submitted to Lesley University.
Recovery instruction in Massachusetts is meeting its goal of bringing the lowest 20% of readers up to the average level. An important finding from this study is the ongoing positive effect that Reading Recovery has on the literacy skills of the initially lowest performing students. The results of this study indicate Reading Recovery children are as likely to pass the third-grade MCAS exam as is the general student population. In 2002, 92.74% of all children in the 47 school districts under observation passed the exam. The passing rate for all Reading Recovery children in the sample obtained for the 1999–2000 Reading Recovery cohort in Massachusetts was slightly higher than this figure, at 92.86%. Reading Recovery children who successfully completed their series of lessons had a passing rate of 95.49%. Reading Recovery claims to bring children up to the average reading level of their classrooms, and these results suggest that the positive benefits of Reading Recovery instruction are sustained 2 years after its conclusion.

On average, children who received any portion of the Reading Recovery intervention also do not show up in the Warning category at any different rate than the overall population. Given that Reading Recovery targets the lowest 20% of readers, the lack of statistical difference suggests Reading Recovery has largely done in Massachusetts what it claims to do. Overall, the findings from this study suggest Reading Recovery is helping to “lift the floor” for the lowest-performing first graders and help them maintain their literacy gains through the end of third grade. Reading Recovery does work, with rates of passing among former Reading Recovery students similar to or even higher than the general population.

Future Research
The findings from this study are quite positive for Reading Recovery, but they should be interpreted with caution given the limitations of the data and methods outlined earlier. The data limitations of this study, however, are generally biased against Reading Recovery, and better data collection with more complete demographic information allowing one to control for relevant factors might very well reveal an even more positive picture of Reading Recovery and its effectiveness.

Reading Recovery is a 12- to 20-week intervention for the lowest 20% of readers in a class. Other contextual factors, such as the classroom or home environments in which children find themselves, may limit or extend the positive effects of Reading Recovery. These factors likely become more and more dominant as the child gets further from first grade. At a minimum, the child has had 2 years between the conclusion of Reading Recovery and sitting for the MCAS reading exam. Although Reading Recovery's goal is to help children construct a complex processing system in reading and writing which allows them to continue learning on their own (known as a self-extending system) without parental involvement or specialty support, the significant time lapse between the conclusion of Reading Recovery and sitting for the exam may mean that these contextual factors matter more as children get further away from the intervention.

When conducting sustained effects studies of Reading Recovery instruction, data at the household and school levels should also be collected and analyzed.

Interventions such as Reading Recovery must be able to demonstrate
their effectiveness. More treatment evaluation based on better data needs to be done. For their evaluation of Reading Recovery’s effectiveness, school districts should collect data on both Reading Recovery and non-Reading Recovery students on an ongoing and well-planned basis in order to allow for individual level comparisons within school districts. Demographic data such as age, race, and poverty status, as well as contextual factors including school size, are essential pieces of data to collect. Controlling for such factors is exceedingly important to determine the effectiveness of Reading Recovery as an intervention, given the known predictive power of these other characteristics.

References


About the Author

Catherine Simpson Bueker received a PhD in sociology in 2003 from Brown University. For 2 years, she did research and program evaluation for Lesley University in Cambridge, MA, and is now teaching sociology at Lasell College in Newton, MA.