

2005–06 National Data Preview: Measuring the Impact of Reading Recovery

Francisco Gómez-Bellengé, National Data Evaluation Center

Note: A more complete statistical summary can be found in the publications page of the National Data Evaluation Center (NDEC) website (www.ndec.us) under the heading “2005–2006 Reading Recovery and Descubriendo la Lectura National Statistical Abstract.”

Summary of Reading Recovery Outcomes

In 2005–2006, 107,928 Reading Recovery students were taught by 12,801 teachers. These were located in 7,543 schools in 2,404 school districts and received professional development from 555 teacher leaders based in 448 different teacher training sites. These in turn were supported by 39 university faculty located in 22 university training centers (Table 1).

The demographics of Reading Recovery students were distinct from

those of the schools from which they come. Compared to their classmates, Reading Recovery students are more likely to be boys (58% vs. 50%), receive free or reduced-priced school lunches (59% vs. 44%), and be non-White (42% vs. 35%).

Overall, 59% of all children served successfully discontinued their series of lessons and reached average reading levels; another 19% received a complete series of lessons and were recommended for further instructional support; 15% had interventions cut short by the end of the academic year; 5% moved while being served; and 3% had other outcomes. Looking only at students who had the opportunity to receive a complete series of lessons, three-fourths (76%) were discontinued successfully and met grade-level expectations.

Summary of Descubriendo la Lectura Outcomes

For Descubriendo La Lectura (DLL), 1,601 students were taught by 222 teachers in 186 schools located in 51 school districts in seven states. They received support from 54 teacher leaders (Table 2).

In terms of demographics, 59% of students taught in DLL were boys; 97% received free or reduced-price school lunches, 100% were Hispanic, and 7% had some disability (most often speech and language). Status outcomes were as follows: 54% of all served successfully discontinued their series of lessons and reached average-reading levels, another 19% received

a complete series of lessons and were recommended for further instructional support, 14% had interventions cut short by the end of the academic year, 7% moved while being served, and 6% had other outcomes. Looking only at students who had the opportunity to receive a complete series of lessons, three-fourths (76%) had their lessons successfully discontinued.

Classroom teacher perceptions of students taught in DLL shifted considerably: In fall they thought that 85% of the students were low readers, compared to only 23% of the students in spring. In other words, 62% of the DLL students moved out of the low-reader classification.

By the end of first grade, 1% of DLL students were classified as learning disabled for reading, compared to 1% for the general Spanish bilingual student population. DLL students were no more likely to be identified as learning disabled for reading than

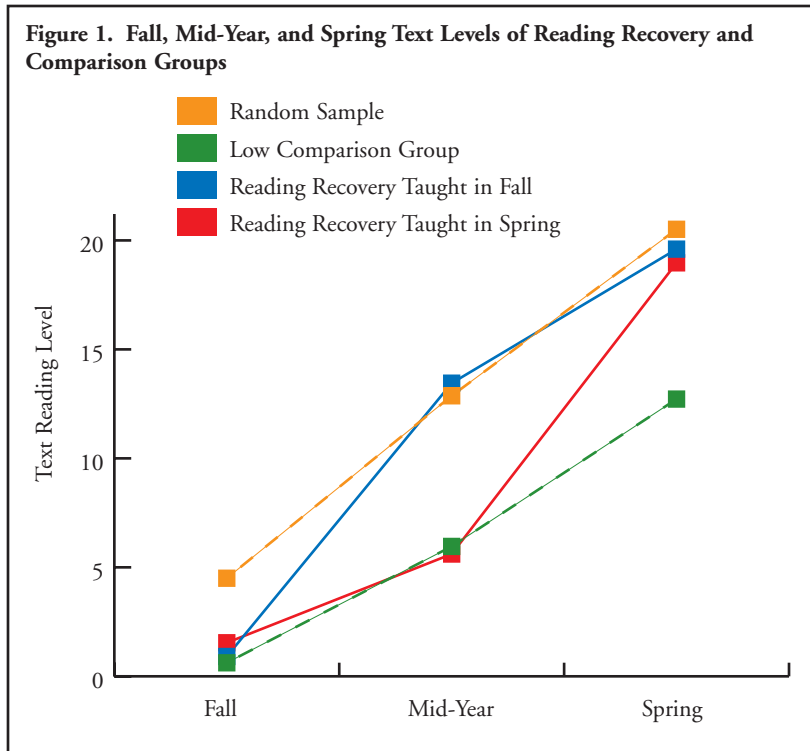
Table 1. Participation in Reading Recovery in the United States 2005–2006

Entity	n
University Training Centers	22
Teacher Training Sites	448
States and Federal Entities*	51
Systems	2,404
Buildings	7,543
Teacher Leaders	555
Teachers	12,801
Reading Recovery Students	107,928
Random Sample for RR	14,253

*including Bureau of Indian Affairs, Department of Defense Domestic, and Department of Defense Overseas

Table 2. Participation in Descubriendo la Lectura in the United States 2005–2006

Entity	n
University Training Centers	7
Teacher Training Sites	43
States	7
Systems	51
Buildings	186
Teacher Leaders	54
Teachers	222
DLL Students	1,601
Random Sample for DLL	327



the general population. This is remarkable because all DLL students were initially at risk.

Comparison Groups

About half of the interventions begin in fall and another half begin at mid-year, with a few interventions beginning in spring. Reading Recovery teachers always select the lowest readers first, so students whose interventions begin in fall have lower initial scores than those who begin at mid-year. As part of its evaluation methodology, the outcomes for Reading Recovery children are compared with two comparison groups in the fall, at mid-year, and at year-end.

Random Sample

Reading Recovery tests a random sample of students in fall, at mid-year, and at year-end. Two students from each Reading Recovery school are included in the national *compari-*

son group. The progress of Reading Recovery random sample children is compared with this group.

Low Comparison Group

Using a statistical technique called *principal component analysis*, National Data Evaluation Center researchers created a single reading score for the comparison group from the six tasks of the Observation Survey. Students were then ranked, with the 15% lowest separated out into a *low comparison group*. We chose the 15% lowest because that is the approximate proportion of students served by Reading Recovery in U.S. schools. This group of students not served by Reading Recovery, but who were equally low readers in fall of Grade 1, can help us see what would happen to Reading Recovery students if they did not benefit from Reading Recovery. The results are illustrated in Figure 1.

In the fall, the comparison group, representing the overall population of first-grade students in schools with Reading Recovery, averaged a Text Reading Level of 4. Reading Recovery students served in fall and the low comparison group averaged a text level of less than 1. Reading Recovery students taught in the second half of the year averaged a little over 1. From fall to mid-year, students who received the Reading Recovery intervention closed the gap. The low comparison group, which did not, fell further behind. Both the yet-to-be-served Reading Recovery students and the low comparison group were six to seven text levels below the comparison group at mid-year.

By spring, the Reading Recovery students served starting at mid-year had nearly caught up to the comparison group, averaging a text level of 18 as compared to 20 for the comparison group. Students served in the first half of the year who did not receive any further support for the second half of the year had nearly identical scores to the comparison group. The low comparison group, however, had an average text level of less than 13, seven text levels less than the comparison group, six less than the first group of Reading Recovery students, and five less than the second group.

For a second-grade teacher receiving a group of students averaging a text level of 20, working with former Reading Recovery students averaging text levels of 18 or 19 is hardly a challenge. These students are not practically distinguishable from average students. Working with students reading below text level 13, however, would challenge the teacher.

Ways to Measure the Impact of Reading Recovery

The impact of an intervention such as Reading Recovery can be measured in different ways.

Statistical significance. At year-end, there was a statistically significant difference between Reading Recovery students and low comparison group students who were low readers in the fall but were not taught in Reading Recovery. Although seemingly very scientific, stating that the difference between the Reading Recovery group

and the low comparison group is statistically significant only means that the difference between two groups is not due to random chance. The problem is that even small differences are statistically significant when comparing large groups. This is not very useful.

Effect size tells us how large the difference between two groups is, regardless of the size of the groups. An effect size of less than 0.2 is considered small; between 0.2 and 0.8, it is medium, and over 0.8, it is large¹.

When we compare Reading Recovery students taught in the first half of the year to the low comparison group at mid-year, the low comparison group averaged a text level of 6, while the Reading Recovery group averaged 13. The computed effect size is 2, which is very large. When comparing Reading Recovery students taught in the second half of the year with the low comparison group, the low comparison group averaged a text level of 13, while the Reading Recovery group taught in the second half of the

Comparing the Cost Effectiveness of Reading Recovery to Title I

To compare the cost of Reading Recovery to other forms of instruction, we will apply the formula used by Leech and Onwuegbuzie (2004).

Let us assume an average teacher salary of \$48,000 and an average effect size of 0.15 for Title I instruction (Borman & D'Agostino, 2006). We assume benefits of 30% for a total of \$62,400. We will assume that Title I teachers work with six groups of 5 students every day for a total of 30 students per year. We will not include the cost of materials since students will use some materials regardless of what intervention they receive. The average cost of Title I instruction per student is \$62,400 divided by 30 students = \$2,080. With an average effect size of 0.15, the cost per effect size is $\$2,080 / 0.15 = \$13,866$ per standardized point mean difference. This is a value of economic significance. It is not meaningful by itself but instead is used to compare interventions. It is the cost per unit of instructional benefit for the student.

For Reading Recovery, we assume the same teacher salary and benefits. Since it is a half-time position and Reading Recovery teachers work with an average of 8.4 Reading Recovery students per year, the full time equivalent is 16.8 Reading Recovery students per year. The cost per student per year is \$3,714 ($\$62,400 / 16.8$ students).

We will make two comparisons: One includes all students served by Reading Recovery, even if for only one

session. The other includes all students who received a complete intervention. The Text Reading Level measure is used to compute effect size.

Comparing all students served by Reading Recovery, even if for only one session, to the low comparison group, we find an effect size of 0.45. Divided by the cost per student ($\$3,714 / 0.45$), this comes to an \$8,253 standardized point mean difference. If we compare students who receive a complete intervention only, the effect size is 0.67 and the standardized point mean difference is \$5,543 ($\$3,714 / 0.65$). Either way, Reading Recovery is substantially less costly than Title I instruction when educational outcomes are taken into account.

This formula can be applied easily by school districts. The bottom line is that when comparing the cost of interventions, it is essential to also compare their instructional outcomes.

References

- Borman, G. D. & D'Agostino, J. V. (2006). Title I and student achievement: A meta-analysis of federal evaluation results. *Educational Evaluation and Policy Analysis, 18*, 309–326.
- Leech, N. L. & Onwuegbuzie, A. J. (2004). A proposed fourth measure of significance: The role of economic significance in educational research. *Evaluation and Research in Education, 18*, 179–198.

¹Effect size is computed using means and standard deviations from fall and spring with the online calculator available at <http://web.uccs.edu/lbecker/Psy590/escalc3.htm>

year averaged 19 at year-end. The computed effect size is 1.1, which is large (see Figure 1). Any effect size over 0.8 is considered large, so the impact of Reading Recovery for these low readers is in fact quite large. This can also be called *practical significance*.²

Pedagogical significance is the impact of the intervention on teaching and learning. The difference between students taught by Reading Recovery teachers in the second half of the year and the comparison group (representing the average for all first graders), two text levels, is relatively small. It is within the normal range of variation in a classroom and can be accommodated with regular classroom instruction. The difference between a group averaging a text level of 20 (average of all first graders) and another 13 (low comparison group never receiving Reading Recovery), on the other hand, requires major accommodations on the part of the teacher. These students will not be able to

read the same materials and may need differentiated instruction even in nonliteracy content areas.

Economic significance can be computed by dividing the cost of the intervention by the effect size. This yields a cost-effectiveness *educational significance index*. (See sidebar article and Leech & Onwuegbuzie, 2004.)

Whether it be the recent research, the new report from the What Works Clearinghouse or the latest evaluation data, evidence continues to accumulate that Reading Recovery is an effective intervention.

References

Leech, N. L. & Onwuegbuzie, A. J. (2004). A proposed fourth measure of significance: The role of economic significance in Educational Research. *Evaluation and Research in Education*, 18, 179–198.

Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology*, 67, 332–339.

² This is also called *clinical significance*. See Kazden, A.E., 1999.

About the Author



Francisco Gómez-Bellengé holds degrees in anthropology from the University of Arizona and Case Western Reserve University, with Ph.D. coursework in education at Cleveland State University. His background includes experience in both research and education. He is involved in a variety of research projects with university faculty and the NDEC staff.

NEW

Native American Readers



Unique set of leveled readers featuring Native American children and their families.



Also Available in French

To order or for more information visit our website at:
www.eaglecrestbooks.com
 or contact us at: info@eaglecrestbooks.com
 phone: (250) 748-3744 • fax: (250) 748-3777