Reading Recovery®

Program description

Reading Recovery® is a short-term tutoring intervention program intended to serve the lowest achieving (bottom 20%) first-grade students. According to the Reading Recovery® web site, lessons incorporate the program’s ten principles: phonological awareness, visual perception of letters, word recognition, phonics/decoding skills, phonics/structural analysis, fluency/automaticity, comprehension, a balanced literacy approach, early intervention, and individual tutoring. Students are chosen for Reading Recovery® by school staff, and selection is based on prior reading achievement, diagnostic testing (the Clay Observation Survey of Early Literacy Achievement), and teacher recommendations. The goals of Reading Recovery® are to promote literacy skills and reduce the number of first-grade students who are struggling to read. The program supplements classroom teaching with one-on-one tutoring sessions, generally conducted as pull-out sessions during the school day. Tutoring, which is conducted by trained Reading Recovery® teachers, takes place daily for 30 minutes over 12–20 weeks.1

Research

Four studies of Reading Recovery® met the What Works Clearinghouse (WWC) evidence standards, and one study met WWC evidence standards with reservations. These five studies included about 700 first-grade students attending elementary schools in diverse settings across the United States. All studies focused on low-achieving students who received the Reading Recovery® intervention in first grade. Generally, outcomes at the end of first grade were used by the WWC to calculate a rating of effectiveness.2 In one study, longer range effects were included.3

Effectiveness

Reading Recovery® was found to have positive effects on students’ alphabets skills and general reading achievement outcomes. The program was found to have potentially positive effects on comprehension and fluency.

<table>
<thead>
<tr>
<th></th>
<th>Alphabets</th>
<th>Fluency</th>
<th>Comprehension</th>
<th>General reading achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating of effectiveness</td>
<td>Positive</td>
<td>Potentially positive</td>
<td>Potentially positive</td>
<td>Positive</td>
</tr>
</tbody>
</table>

1. The WWC does not verify the accuracy of the developer’s description of the intervention.
2. The evidence presented in this report is based on available research. Findings and conclusions may change as new research becomes available.
3. Additional findings on outcomes measured at later time points are shown in Appendix A.4.4.
Developer and contact

Scope of use
Reading Recovery® was developed in the mid-1970s by Dr. Clay, who first tested the program in New Zealand. According to the RRCNA, more than 1.5 million first graders in 48 states and the Department of Defense Dependents Schools have been served in the United States since Reading Recovery® was introduced in 1984. Reading Recovery® is also used in New Zealand, Australia, Canada, and the United Kingdom.

Teaching
According to the Reading Recovery® web site, lessons incorporate the program’s ten principles: phonological awareness, visual perception of letters, word recognition, phonics/decoding skills, phonics/structural analysis, fluency/automaticity, comprehension, a balanced literacy approach, early intervention, and individual tutoring. Each Reading Recovery® lesson consists of reading familiar and novel stories, manipulating letters and words, and writing and assembling stories. Lessons are interactive between teacher and student, with the teacher carefully monitoring each child’s reading behavior. Reading Recovery® lessons are discontinued when children demonstrate the ability to consistently read at the average level for their grade—between weeks 12 and 20 of the program. Those who make progress but do not reach average classroom performance after 20 weeks are referred for further evaluation and a plan for future action. Teacher training includes a one-year, university-based training program.

Cost
Reading Recovery® is available on a nonprofit, no royalty basis. Costs for the program involve start-up costs and ongoing costs. To establish a Reading Recovery® site—a district or group of districts representing multiple schools—a teacher leader must first be trained. This start-up cost includes paying salary, paying university tuition for the Reading Recovery® coursework, and covering the costs of books and materials. Sites must also build a one-way mirror and sound system to monitor training for the teachers. In addition to salary, travel, and program support costs for the teacher leader, costs for teachers include paying salaries and benefits for the time they dedicate to Reading Recovery® and paying tuition for training. Books and materials for lessons and evaluation as well as ongoing professional development for both teacher leaders and teachers should also be figured into the costs.

In addition to the teacher training described above, in 2006 the cost of program materials was approximately $100 per student served (calculated by the RRCNA as an average over the past five years, 2002–06). Sites pay an annual data evaluation fee of $250 a site plus $3.50 per student served. Sites implementing the program also pay annual technical support fees, which vary by the university that provides the Reading Recovery® training. Because of the cost and staff needed for the intervention, a typical school with one Reading Recovery® teacher will serve 4 or 5 students a semester.

---

### Improvement index

<table>
<thead>
<tr>
<th>Improvement index</th>
<th>Alphabets</th>
<th>Fluency</th>
<th>Comprehension</th>
<th>General reading achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Average: +34 percentile points</td>
<td>Average: +46 percentile points</td>
<td>Average: +14 percentile points</td>
<td>Average: +32 percentile points</td>
</tr>
<tr>
<td>Range: −10 to +50 percentile points</td>
<td>Range: +32 to +49 percentile points</td>
<td>Range: +6 to +21 percentile points</td>
<td>Range: −5 to +50 percentile points</td>
<td></td>
</tr>
</tbody>
</table>

---

4. These numbers show the average and range of improvement indices for all findings across the studies.
The WWC reviewed 78 studies that examined the effects of Reading Recovery®. Four studies (Baenen, Bernhole, Dulaney, & Banks, 1997; Pinnell, DeFord, & Lyons, 1988; Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994; and Schwartz, 2005) met WWC evidence standards. One study (Iverson & Tunmer, 1993) met WWC standards with reservations. The remaining 73 studies did not meet WWC evidence screens.

**Met evidence standards**

Baenen et al. (1997) was a randomized controlled trial that focused on first-grade students from Wake County, North Carolina. The WWC review focuses on the outcomes of students who qualified for and were randomly assigned to either the Reading Recovery® intervention or a comparison group. From an original sample size of 168, outcomes were assessed at three time points: end of first grade (n = 147), end of second grade (n = 147), and end of third grade (n = 127). Although the WWC used only the results at the end of first grade to determine the intervention rating, information on the additional findings can be found in Appendix A4.4.

Pinnell, DeFord, and Lyons (1988) was a randomized controlled trial. The study sample was first-grade students distributed across 14 schools in Columbus, Ohio. Two groups were formed by randomly assigning students to an intervention group, which received Reading Recovery® in addition to their regular classroom instruction (n = 38), or to a control group, which received an alternate compensatory program (n = 53). This comparison met WWC evidence standards. Pinnell et al. (1994) was a randomized controlled trial that randomly assigned 10 low-achieving first-grade students in each of 10 Ohio schools. The WWC review focuses only on the eight schools that successfully implemented randomization for the intervention (n = 31) and comparison (n = 48) conditions.

Schwartz (2005) was a randomized controlled trial of first-grade students from 14 states. The WWC focused on the 37 students across several schools who were randomly assigned to receive the intervention during the first half of the year. The other 37 students, who were randomly assigned to receive the intervention during the second half of the year, served as the comparison group during the first half of the year. The groups were compared at mid-year, before the comparison group had begun receiving Reading Recovery®.

**Met evidence standards with reservations**

Iverson and Tunmer (1993) was a quasi-experimental design study that included first-grade students from 30 school districts in Rhode Island. The study compared outcomes for students participating in Reading Recovery® (n = 32) with students in a comparison group who did not receive Reading Recovery® (n = 32), who were matched on the basis of pretest scores. The comparison group received standard small group, out-of-class support services.

---

5. Because Reading Recovery® is designed to improve the reading skills of low-achieving first-grade readers, the appropriate comparison groups for determining the intervention’s effectiveness are similar low-achieving first-grade readers who did not receive Reading Recovery®. Many of the studies screened did not meet evidence standards because they used inappropriate comparison groups, such as higher achieving first-grade readers, to draw conclusions about the effectiveness of the program.

6. A third group of students qualified for and received Reading Recovery® outside of regular classroom instruction, but were also taught by a Reading Recovery®–trained teacher when they were in their regular classroom (n = 96). Although this comparison met evidence standards with reservations, it was not considered in the intervention rating because it went beyond the standard delivery of the program. However, results are reported in Appendices A4.1–A4.3.

7. Although the original study included analyses of additional interventions implemented at additional schools, only the schools that randomly assigned students to Reading Recovery® or the comparison group were relevant to this review. For more details about the original study, see Appendix A1.3.

8. Assessments were also made at the end of the year, but they were not appropriate for the WWC’s analysis because by then both groups of low-achieving students had received the intervention. Additional comparison groups of low-average and high-average readers were not used by the WWC because these students were not eligible for Reading Recovery®.

9. The study also included a third group of students (n = 32) who used a modified version of Reading Recovery®, which provided explicit instruction in letter-phoneme patterns instead of the letter identification segment. This group was also compared with the comparison group. Although this comparison met evidence standards with reservations, it was not considered in the intervention rating because it went beyond the standard delivery of the program. However, results are reported in Appendices A4.1 and A4.3.
Findings
The WWC review of interventions for beginning reading addresses student outcomes in four domains: alphabetics, reading fluency, comprehension, and general reading achievement. Reading Recovery® studies included in this report cover all four domains and most of the constructs within each domain. The findings below present the authors’ estimates and WWC-calculated estimates of the size and the statistical significance of the effects of Reading Recovery® on students. The results are presented by domain and construct for all the Reading Recovery® studies that the WWC reviewed.

For the four beginning reading domains, subtests of the Clay Observation Survey were used in some of the studies. The Clay Observation Survey was developed by Dr. Marie Clay, who also developed Reading Recovery®.

Alphabetics. Two studies examined the effects of Reading Recovery® on the phonemic awareness construct. Schwartz (2005) reported no statistically significant effects for the phonemic awareness measures—the deletion task and the Yopp-Singer Phoneme Segmentation Test—but the effects on both measures were positive and considered substantively important based on the WWC criteria (that is, at least 0.25). Iverson and Tunmer (1993) reported, and the WWC confirmed, statistically significant positive effects of the Reading Recovery® intervention on two phonemic awareness measures—a phone deletion task and the Yopp-Singer Phoneme Segmentation Test.

Three studies examined the effects of Reading Recovery® on the print awareness construct in the alphabetics domain. Pinnell, DeFord, and Lyons (1988) did not find a statistically significant effect for Reading Recovery® on the Letter Identification subtest of the Observation Survey. Schwartz (2005) found, and the WWC confirmed, statistically significant positive effects on the Concepts about Print subtest of the Observation Survey. Iverson and Tunmer (1993) found a statistically significant positive effect of Reading Recovery® on the Concepts about Print subtest of the Observation Survey. The significance of the effect was confirmed by the WWC.

Overall, in the alphabetics domain, two studies with strong designs met WWC evidence standards and demonstrated statistically significant positive effects. One additional study met WWC evidence standards with reservations and showed statistically significant positive effects.

Effectiveness

10. For definitions of the domains, see the Beginning Reading Protocol.
11. In this case, the author did not control for pretest differences between groups; however, the WWC did account for pretest differences.
**Effectiveness (continued)**

**Fluency.** Schwartz (2005) found, and the WWC confirmed, positive and statistically significant effects of *Reading Recovery*® on the Slosson Oral Reading Test–Revised and the Text Reading Level subtest of the Observation Survey. In the fluency domain, there was one study with a strong design that demonstrated statistically significant positive effects.

**Comprehension.** Two studies examined the effects of *Reading Recovery*® on the reading comprehension construct. Pinnell, DeFord, and Lyons (1988) found a positive and statistically significant effect of *Reading Recovery*® on the Reading Comprehension subtest of the Comprehensive Test of Basic Skills (CTBS). The significance of the effect was confirmed by the WWC. Schwartz (2005) reported no statistically significant effect of *Reading Recovery*® on the Degrees of Reading Power Test.

One study examined the effect of *Reading Recovery*® on the vocabulary construct of the comprehension domain. Pinnell, DeFord, and Lyons (1988) found, and the WWC confirmed, a positive and statistically significant effect of *Reading Recovery*® on the Reading Vocabulary subtest of the CTBS.

In the comprehension domain, there were two studies with strong designs. One study showed statistically significant positive effects, and the other study showed an indeterminate effect.

**General reading achievement.** Baenen et al. (1997) did not find a statistically significant effect of *Reading Recovery*® on grade retention. Pinnell, DeFord, and Lyons (1988) found, and the WWC confirmed, positive and statistically significant effects of *Reading Recovery*® on two subtests of the Observation Survey: Hearing and Recording Sounds in Words (Dictation) and Writing Vocabulary. Pinnell et al. (1994) found statistically significant positive effects of *Reading Recovery*® on the Gates-MacGinitie, the Dictation subtest of the Observation Survey, and the Woodcock Reading Mastery Test–Revised. The statistical significance of the effects was confirmed by the WWC. Schwartz (2005) found, and the WWC confirmed, positive and statistically significant effects of *Reading Recovery*® on two subtests of the Observation Survey: Dictation and Writing Vocabulary.

In the general reading achievement domain, there were three studies with strong designs and statistically significant positive effects. One study had a strong design with indeterminate effects. One additional study met WWC evidence standards with reservations and demonstrated statistically significant positive effects.

**Rating of effectiveness**

The WWC rates the effects of an intervention in a given outcome domain as positive, potentially positive, mixed, no discernible effects, potentially negative, or negative. The rating of effectiveness takes into account four factors: the quality of the research design, the statistical significance of the findings, the size of the difference between participants in the intervention and the comparison conditions, and the consistency in findings across studies (see the [WWC Intervention Rating Scheme](#)).

---

12. The level of statistical significance was reported by the study authors or, where necessary, calculated by the WWC to correct for clustering within classrooms or schools and for multiple comparisons. For an explanation, see the [WWC Tutorial on Mismatch](#). See [Technical Details of WWC-Conducted Computations](#) for the formulas the WWC used to calculate the statistical significance. In this report, such adjustments were made for Pinnell, DeFord, and Lyons (1988); Iverson and Tunmer (1993); and Pinnell et al. (1994).
Improvement index

The WWC computes an improvement index for each individual finding. In addition, within each outcome domain, the WWC computes an average improvement index for each study and an average improvement index across studies (see Technical Details of WWC-Conducted Computations). The improvement index represents the difference between the percentile rank of the average student in the intervention condition versus the percentile rank of the average student in the comparison condition. Unlike the rating of effectiveness, the improvement index is entirely based on the size of the effect, regardless of the statistical significance of the effect, the study design, or the analyses. The improvement index can take on values between –50 and +50, with positive numbers denoting results favorable to the intervention group.

The average improvement index for alphabetics is +34 percentile points across three studies, with a range of –10 to +50 percentile points. For fluency, the average improvement index is +46 percentile points, with a range of +32 to +49 percentile points across outcomes in one study. For comprehension, the average improvement index is +14 percentile points across two studies, with a range of +6 to +21 percentile points. For the general reading domain, the average improvement index was +32 percentile points across five studies, with a range of −5 to +50 percentile points.

Summary

The WWC reviewed 78 studies that investigated the effects of Reading Recovery®. Four studies met WWC evidence standards and one met WWC evidence standards with reservations. Based on these five studies, the WWC found Reading Recovery® to have positive effects in the alphabetics and general reading achievement domains. Reading Recovery® was found to have potentially positive effects in the fluency and comprehension domains. The evidence presented in this report may change as new research emerges.

References

Met WWC evidence standards


Additional sources:


Additional sources:


References (continued)


Met WWC evidence standards with reservations


Additional source:


Did not meet WWC evidence screens


Collins, E. W. (2000). The immediate and sustained effects of the Reading Recovery program on grade one and grade four


**Additional source:**


References (continued)


References (continued)

V. Hoffman, D. L. Schallert, C. M. Fairbanks, J. Worthy, & B. Maloch (Eds.), *Fiftieth yearbook of the National Reading Conference* (pp. 519–532). Chicago: National Reading Conference.23


**Additional source:**


References (continued)  


For more information about specific studies and WWC calculations, please see the WWC Reading Recovery® Technical Appendices.

13. The sample is not appropriate to this review: the parameters for this WWC review specified that students should be in grades kindergarten through 3; this study did not disaggregate students in the eligible range from those outside the range.
14. Does not use a strong causal design: this study was a quasi-experimental design but did not use achievement pretests to establish that the comparison group was equivalent to the intervention group at baseline.
15. Incomparable groups: this study was a quasi-experimental design that used achievement pretests but it did not establish that the comparison group was comparable to the treatment group prior to the start of the intervention.
16. Does not use a strong causal design: this study did not use a comparison group.
17. Does not use a strong causal design: in this study, which used a quasi-experimental design, data were presented only for students who successfully completed the intervention, which does not provide a direct test of the intervention as a whole.
18. Does not use a strong causal design: this study, which used a quasi-experimental design, had a confounding factor. The Reading Recovery intervention was used without proper Reading Recovery materials, and the instructors had not been fully trained. This makes it difficult to attribute study outcomes to Reading Recovery.
19. The sample is not appropriate to this review: this study did not focus on students learning to read in English, one of the parameters for this WWC review.
20. Does not use a strong causal design: this is a qualitative study.
21. Complete data are not reported: the WWC could not evaluate the design because complete data were not reported. Attempts to contact the authors for more information were unsuccessful.
22. High overall attrition: this study reported an attrition rate of greater than 20 percentage points.
23. The sample is not appropriate to this review: the parameters for this WWC review specified student outcome measures, but this study did not focus on students.
24. Disruption: this study, which used a quasi-experimental design, exhibited disruption problems that made it difficult to attribute study outcomes to the intervention, as delivered.
25. Does not use a strong causal design: this study, which used a quasi-experimental design, tested only a portion of the Reading Recovery curriculum, making it difficult to attribute study outcomes to Reading Recovery.

26. Does not use a strong causal design: for the portion of the sample of interest for this WWC review, there was one of two issues. Either there was a confound, with the Reading Recovery intervention being modified or combined with other interventions, making it difficult to attribute study outcomes to the intervention. Or the study did not establish that the comparison group was comparable to the treatment group prior to the start of the intervention.

27. The outcome measures are not relevant to this review: the outcomes in this study did not address one of the domains of interest in this review.