

SECTION I

READING RECOVERY IS HIGHLY SUCCESSFUL WITH THE LOWEST-PERFORMING FIRST-GRADE STUDENTS.

Of all charges leveled against Reading Recovery, the question of effectiveness is most serious. The Internet letter authors limited the kind of evidence examined and excluded research studies that did not support their claims. If a research study cited in the Internet letter included evidence that balanced criticisms with favorable findings for Reading Recovery, the positive findings were not reported. In addition, the Internet letter ignored large amounts of evaluation data supporting Reading Recovery.

This section

- A. provides a more complete review of research on Reading Recovery's effectiveness and subsequent literacy gains,
- B. describes Reading Recovery's internal evaluation system, and
- C. analyzes biases in the evidence cited by the Internet letter.

The information presented here represents a broader view of the research such as that recently outlined by Michael Pressley in his Oscar Causey Research Address to the National Reading Conference (2001).

A. Research demonstrates Reading Recovery effectiveness.

Substantial evidence documents Reading Recovery's success with the lowest-performing first-grade students in a wide range of educational settings. This evidence is provided by a variety of study designs, including experimental and control group studies both by researchers associated with Reading Recovery and by independent researchers.

1. Research in peer-reviewed journals

The following research published in peer-reviewed journals documents Reading Recovery's effectiveness with first-grade children.

Substantial evidence documents Reading Recovery's success with the lowest-performing first-grade students in a wide range of educational settings.

Programs with the most comprehensive models of reading—the most complete instructional interventions—have greater impact than programs addressing only a few components of the reading process, and Reading Recovery and Success for All include several components.

(Finding of a 1993 study by Wasik and Slavin)

a. **Wasik and Slavin (1993)** considered the effectiveness of five tutorial programs from two perspectives: empirical and pragmatic. The authors reviewed quantitative and qualitative research on Reading Recovery, Success for All, Prevention of Learning Disabilities, Wallach Tutoring Program, and Programmed Tutorial Reading. The authors' general conclusions across programs were

- Programs with the most comprehensive models of reading—the most complete instructional interventions—have greater impact than programs addressing only a few components of the reading process, and Reading Recovery and Success for All include several components.
- Using tutors is not enough; the content of the program and the instructional delivery may be important variables.
- Using certified teachers obtains substantially better results than using paraprofessionals.

The authors' specific conclusions about Reading Recovery included

- Reading Recovery brings the learning of many of the lowest-achieving students up to average-achieving peers.
- Effects of Reading Recovery are impressive at the end of the implementation year, and effects are maintained for at least two years.
- Evaluation results on lasting effects are positive but complex.
- Only Reading Recovery has attempted to assess implementation and its effect on outcome data.

Although the authors raised some methodological issues about Reading Recovery research and about students served, they concluded that the rapidly expanding use of Reading Recovery throughout the United States shows that the program is practical to use.

b. **Center, Wheldall, Freeman, Outhred, and McNaught (1995)** evaluated the effectiveness of Reading Recovery schools in New South Wales. Low-achieving children were randomly assigned to either Reading Recovery (n=31) or a control group (n=39) of low-progress students who had not entered Reading Recovery by November. A third group (n=39) consisted of students from five matched schools. By the end of the study, sample sizes were 23, 16, and 32 respectively. Measures used were Clay's Diagnostic Survey, Burt Word Reading Test, Neale Analysis of Reading Ability, Waddington Diagnostic Spelling Test,

Phonemic Awareness Test, Cloze Test, Word Attack Skills Test, and Woodcock Reading Mastery.

At post-test evaluation (15 weeks after the pre-test) an independent assessment showed that Reading Recovery students scored significantly higher on all tests measuring reading in context and in isolation. Of the eight measures reported, the only ones that did not differ significantly were a cloze test and a phonemic awareness measure. At short-term maintenance (15 weeks after the post-test) the Reading Recovery control group still scored significantly higher than the control group on six of the eight measures, including Clay's text reading measure and several standardized measures of text and word reading. At this point the Reading Recovery group also scored significantly higher than the control group on phonemic awareness.

The study's published results for medium-term maintenance (12 months after the post-test) appear to have errors. The authors report "no overall significant group effect, $F(8,30) = 0.262$, $p = .0268$ " (p. 253). There appear to be several typos and errors in this statistical statement beyond the inclusion of an additional closing parenthesis. An F value of 2.62 would match the probability level of .0268. Since the authors state that "significant multivariate results ($\alpha = 0.05$) were followed up by univariate pairwise multiple comparisons ($\alpha = 0.01$)" (p. 250), the conclusion should be that the MANOVA revealed an overall significant group effect in favor of Reading Recovery. Still, the only univariate result was for text reading. The authors point out that the reduced difference between the Reading Recovery and control groups found in the 12-month follow-up could be due to the fact that 15 of the 31 control group students (probably those with the lowest scores) had been eliminated from the control group to receive Reading Recovery instruction.

The study provides strong, independent replication of the pattern of results found in other research and in the U.S. national evaluation data for all participating students. The authors state that their "results clearly indicate that low-progress students, exposed to 15 weeks of Reading Recovery, outperformed control students on Clay book-level and Burt Word Reading tests and on all Set 2 tests which measure reading and writing words in context and isolation" (p. 256). Despite a number of qualifications related to metalinguistic measures, the article reports independently measured and extremely large effect size for text reading, 3.05 and 1.55 for post-test and short-term maintenance respectively (p. 253).

c. **Iversen and Tunmer (1993)** conducted a study to determine whether the Reading Recovery program would be more effective if systematic instruction in phonological recoding skills were incorporated into the program. Three matched groups of 32 at-risk readers were compared:

- children taught by teachers who received Reading Recovery training,
- children taught by teachers who received Reading Recovery training that included phonological recoding skills as part of the lesson, and
- children who received a standard intervention (not Reading Recovery).

Measures included all six tasks of the Diagnostic Survey, Dolch Word Recognition Test, Yopp-Singer Phoneme Segmentation Test, Phoneme Deletion Test, and Pseudoword Decoding Task.

The critical finding in this study was that the two Reading Recovery groups performed at very similar levels when Reading Recovery lessons were successfully completed (discontinued). Both groups performed much better on all measures than children in the standard intervention group, and they often performed significantly better than classroom controls (especially on phonological segmentation and phoneme deletion). Results revealed that the modified Reading Recovery group reached levels of performance required for discontinuing more quickly than the standard Reading Recovery group. Authors acknowledged that both the standard and modified Reading Recovery programs included explicit instruction in phonological awareness.

d. **Sylva and Hurry (1996)** evaluated the effectiveness of two different interventions (Reading Recovery and Phonological Training). Their study included almost 400 children from seven English local authorities. Although the sample was diverse, inner city children were over-represented. The schools included 22 Reading Recovery schools, 23 Phonological Intervention schools, and 18 control schools. The measures used included the British Ability Scale Word Reading, Neale Analysis of Reading, Clay's Diagnostic Survey (five tasks), Assessment of Phonological Awareness, British Ability Scale Spelling, and background information on each child.

During the intervention year, the effect of the phonological intervention was more specific than Reading Recovery, enhancing children's phonological awareness and influencing their letter identification, dictation, and writing vocab-

ulary, but not their text reading skills. Reading Recovery children made significantly more progress than the control groups on every measure of reading. At the end of the second year, the effects of phonological intervention were still evident in enhanced word reading scores, but there was no effect on comprehension. In comparison, the Reading Recovery children were still six months ahead of the control children on word and text reading.

In their long-term follow-up four years later, Hurry and Sylva (1998) concluded that Reading Recovery was still effective because almost 70% of the children who had received Reading Recovery were still within the average band of their class in Grade 6, while only 55% of those who received the phonological intervention were within the average class band. Reading Recovery was particularly effective at helping children who were most socially disadvantaged and who were the weakest readers at age 6. While Reading Recovery was more expensive than the phonological intervention in the year of delivery, over the pupils' whole elementary school career it cost only 10% more than the general remedial support provided in control schools. The authors concluded that time-limited intervention is not so expensive in the longer term.

e. Pinnell, Lyons, DeFord, Bryk, and Seltzer's (1994) study systematically compared Reading Recovery to three other instructional models of early intervention. In this study (N=324) the lowest-achieving first-grade students from 40 different schools in 10 different school districts were randomly assigned within schools to one of five groups:

- Reading Recovery,
- a Reading Recovery-like intervention with partially trained teachers,
- a skills-based individual intervention,
- small group instruction offered by Reading Recovery teachers, or
- a control group.

Measures included those used in Reading Recovery as well as generally known reading tests (Gates-MacGinitie Reading Test and Woodcock Reading Mastery). The study employed a formal experimental design that used split plots to control effects that may result from differing cultures among school districts or individual schools. The difficulty of small standard errors in analysis of data at the student level was addressed by using the Hierarchical Linear Model for data analysis. The results of the study

Reading Recovery subjects performed significantly better than any other treatment and comparison group on all measures. Essential differences were related to individual instruction, the lesson framework (combination of techniques), and teacher training.

(Findings of 1994 study by Pinnell, Lyons, DeFord, Bryk, and Seltzer)

“Thus, in answer to the question ‘Does Reading Recovery work?’ we must respond in the affirmative. It is clear that many children leave the program with well-developed reading strategies, including phonemic awareness and knowledge of spelling. Although some initially low-achieving students will succeed without Reading Recovery, evidence indicates that many who would not succeed do so as a result of this intervention.”

(Shanahan and Barr, 1995)

were definitive: Reading Recovery subjects performed significantly better than any other treatment and comparison group on all measures. Essential differences were related to individual instruction, the lesson framework (combination of techniques), and teacher training.

- f. Pinnell (1997) reviewed quantitative and qualitative research studies performed by Reading Recovery practitioners that have demonstrated positive effects of Reading Recovery on reading outcomes (including generally recognized measures), successful replicability in diverse settings, positive effects on home relations, improved teacher behaviors and teacher learning, and maintenance of learning gains over time.
- g. Shanahan and Barr (1995) published a comprehensive and independent evaluation of Reading Recovery. Although the Internet letter cites this evaluation as evidence of the ineffectiveness of Reading Recovery, it in fact offers substantial support for Reading Recovery’s effectiveness.

Evidence firmly supports the conclusion that Reading Recovery does bring the learning of many children up to that of their average-achieving peers. Thus, in answer to the question “Does Reading Recovery work?,” we must respond in the affirmative. It is clear that many children leave the program with well-developed reading strategies, including phonemic awareness and knowledge of spelling. Although some initially low-achieving students will succeed without Reading Recovery, evidence indicates that many who would not succeed do so as a result of this intervention. (p. 989)

That Reading Recovery has been so successful is laudatory. It has proven to be a robust program, both in terms of its consequences for student learning and in replicability across sites. Further, it has been a significant force in shaping the way we view early literacy development. (p. 992)

In summary, many research studies published in peer-reviewed journals document Reading Recovery’s effectiveness with first-grade children. Independent researcher Elfreida Hiebert (1994) wrote that “a high percentage of Reading Recovery tutees can orally read at least a first grade text at the end of Grade 1....Once a program is in place, there appears to be considerable fidelity in the results” (p. 21). This is high praise when one considers that Reading Recovery children are the lowest literacy achievers at the beginning of first grade.

2. Research and evaluation studies on improved gains in later grades

Many research and evaluation studies demonstrate that Reading Recovery students maintain and improve their gains in later grades. The ones listed here have used widely accepted standardized measures or state assessment measures, or both.

- a. **Rowe (1995)**, an Australian researcher, studied the progress made in reading by children from school entry to Grade 6 in Victoria, Australia. The sample included 5,092 students and 256 classes in 92 schools. The researcher's intent was not specifically to study Reading Recovery, but information on Reading Recovery's effectiveness emerged as an outcome. The longitudinal design involved repeated measures nested within classes and schools and repeated measures on schools. The second design involved cross sections of students nested within schools that were changing over time. Rowe used several measures to gather student information: Reading Achievement, Primary Reading Survey Test, Test of Reading Comprehension, English Profile, and Reading Bands.

Rowe found that Reading Recovery children benefited notably from participation in the intervention. Reading Recovery appeared to be meeting its intended purpose for those children involved. By Grades 5 and 6, Reading Recovery students were distributed across the same score range as the general school population, but with fewer low scores. Rowe's analysis provided evidence that Reading Recovery had removed the tail-end of the achievement distribution.

- b. **Brown, Denton, Kelly, and Neal (1999)** used two standardized tests to assess California students' continuing achievement through fifth grade. Researchers measured achievement of 760 students who were served in Reading Recovery between 1993 and 1998. Student performance in second through fifth grades was assessed using the Iowa Tests of Basic Skills and Stanford Achievement Test, Ninth Edition. The authors reported that "more than three-fourths of the children who successfully completed Reading Recovery achieved standardized test scores in the average or above average range" (p. 10). Considering that these Reading Recovery students began at the bottom of their class in first grade, their subsequent progress through fifth grade was impressive.
- c. **Askew, Kaye, Frasier, Mobasher, Anderson, and Rodríguez (2002)** collected longitudinal data on former Reading Recovery children in 45 randomly selected

“A high percentage of Reading Recovery tutees can orally read at least a first grade text at the end of Grade 1....Once a program is in place, there appears to be considerable fidelity in the results.”

(Heibert, 1994)

schools through fourth grade. The study focused on discontinued children (those students who met the rigorous criteria for success) in order to see if children who reached average performance in Grade 1 continued to score within average ranges in subsequent years. At the end of fourth grade, a large majority of these children had scores considered to be average or meeting passing criteria on standardized (Gates-MacGinitie Reading Test) and state assessment measures, a very satisfactory outcome in their school settings. They were generally perceived by their fourth-grade teachers as performing within average ranges of their classrooms. Relatively few were placed in tertiary or remedial settings. Findings match Juel's (1988) conclusions that children who are average readers in Grade 1 remain average readers in Grade 4, supporting the need for intervention in first grade.

- d. **Schmitt and Gregory (2001)** conducted a statewide follow-up study in Indiana to examine subsequent performance of former successful Reading Recovery children in Grades 2, 3, and 4. The study demonstrated that the majority of the 271 randomly selected Reading Recovery children were performing as well as a random sample of 277 of their grade-level peers on the vocabulary and comprehension subtests of the Gates-MacGinitie Reading Test. Similarly, third-grade student results on the Comprehensive Test of Basic Skills-5/Terra Nova Form B reflected a normal distribution with a mean at the 45th percentile for these students who had been the lowest-achieving first graders.
- e. **Pinnell (1989)** evaluated two cohorts of students. The purposes of the study were to explore whether Reading Recovery could succeed with low-achieving children and to determine whether those children maintained their gains. The lowest-achieving children were randomly assigned either to Reading Recovery or to a control group served daily in individual lessons taught by a trained paraprofessional (not a Reading Recovery teacher). Both groups were compared with a random sample of average and high-progress first graders as an indication of average progress. The study used all six tasks of Clay's Diagnostic Survey,² a writing sample, and the Comprehensive Test of Basic Skills (two subtests). Pinnell found that in the full Reading Recovery program, Reading Recovery children scored significantly better than control children on seven of the nine

² Until the Observation Survey was published in 1993, it was known as the Diagnostic Survey (Clay, 1985).

diagnostic measures at the end of first grade. They compared well with the random sample group. Reading Recovery children were followed in second and third grade to determine their performance in text reading in subsequent years. Reading Recovery children remained superior in comparison with the control group.

In addition to these studies, many school districts and teacher training sites conduct their own independent analyses through their research departments. Typically these analyses use the standardized measures that the districts use with all children. One such study (Quay, Steele, Johnson, & Hortman, 2001) in a Georgia school district compared Reading Recovery children with a control group who were equivalent in gender, ethnicity, and achievement. “At the end of the school year, multivariate and univariate analyses of variance indicated that the Reading Recovery children were significantly superior to the control group children on: (a) The Iowa Test of Basic Skills Language Tests; The Gates MacGinitie Reading Test; (c) the six tests of An Observation Survey of Early Literacy Achievement; (d) classroom teachers’ assessments of achievement in mathematics, oral communication, reading comprehension, and written expression; (e) classroom teachers’ ratings of personal and social growth in work habits, following directions, self-confidence, social interaction with adults, and social interaction with peers; and (f) promotion rates” (p. 7). These results are especially significant considering that all teachers in this study were in their initial training year.

3. Studies on improved student self-esteem

Two studies published in refereed journals have revealed that Reading Recovery students experience gains in self-concept.

- a. **Cohen, McDonnell, and Osborn** (1989) studied the impact of Reading Recovery on students’ beliefs regarding their competence and capacity to direct their own learning activities. They used causal attribution (Weiner, 1972) and self-efficacy (Bandura, 1977) to support the theoretical framework. Participants included 138 first graders divided among the following groups: 50 were in Reading Recovery, 48 were in remedial reading groups of five or six students each, and 40 were randomly selected from their higher-achieving classmates.

After the interventions, children were tested on two scales to measure attributions and self-efficacy. Results demonstrated that successful Reading Recovery children had profiles similar to high-achieving students, and they more readily attributed their success in school to ability, effort, and mood than did the students in the remedial groups.

The Reading Recovery students also judged themselves to be more competent on school-related tasks (self-efficacy) than the other low-achieving students. These results support the notion that children have positive self-esteem when they leave Reading Recovery.

b. **Rumbaugh and Brown (2000)** studied the effects of Reading Recovery participation on students' self-concept. The treatment group was comprised of 57 students from nine elementary schools who were selected for Reading Recovery instruction in the first week of school. The 46 students in the control group had diverse reading and writing abilities and were not enrolled in any reading intervention or in special education. The control group came from a single elementary school.

All participants were administered the Joseph Pre-School and Primary Self-Concept Screening Test in early September prior to the treatment and again in mid-December.

There were statistically significant differences between Reading Recovery students and control students on the Global Self-Concept and Significance domain scores. Hence, the authors concluded

- Reading Recovery participation does affect positively students' Global Self-Concept scores.
- The meaningful effect of Reading Recovery participation on students' self-concept is related to the additional attention, or Significance domain, that students receive during several months of Reading Recovery.
- The initial positive effect on students' self-concept cannot be attributed to increased growth in independence or cognitive factors.

Based on their results, Rumbaugh and Brown concluded:

School districts that choose to implement and maintain a Reading Recovery program would reap considerable benefits. One of the systemic advantages could be that the districts gain students who experience improved self-concepts due to enhanced feelings of significance. Not only will the Reading Recovery participants most likely become independent readers, they will also most likely become more confident, positive, self-accepting, proud, adaptable, and eager to complete tasks. (p. 28)

As in any program or in any classroom, children who are failing are likely to have self-esteem difficulties. Yet one of the first reports from both parents and first-grade classroom teachers is about the change in Reading Recovery students' self-esteem when they are making progress in Reading Recovery.

“School districts that choose to implement and maintain a Reading Recovery program would reap considerable benefits.... Not only will the Reading Recovery participants most likely become independent readers, they will also most likely become more confident, positive, self-accepting, proud, adaptable, and eager to complete tasks.”

(Rumbaugh and Brown, 2000)

B. Evaluation evidence supports Reading Recovery effectiveness.

Reading Recovery is very successful with its targeted population, first-grade students having most difficulty with early literacy acquisition—the lowest achievers in the class. This success has been carefully documented for 17 years in data gathered and analyzed on every single U.S. student enrolled in Reading Recovery. All data are reported to the National Data Evaluation Center located at The Ohio State University.

Findings are reported in national, state, and district-level technical reports that are designed to present annual evaluations of systematic, simultaneous replications of Reading Recovery. These reports are widely disseminated to state legislators, state boards of education, local school boards, superintendents, and principals.

Since 1984, Reading Recovery in the United States has collected and analyzed data and reported results for more than one million children. Results confirm Reading Recovery’s success with the lowest-performing first graders, as detailed in the table below.

Since 1984, Reading Recovery in the United States has collected and analyzed data and reported results for more than one million children.

Facts About Reading Recovery Student Success in 2000–2001*

149,009	students served
112,814	received a full series of lessons
86,009	discontinued lessons (successful completion)
36,195	did not receive a full series of lessons

Percentage of Students Who Successfully Completed Lessons

78%	of students who received a full series of lessons successfully met discontinued criteria
59%	of students served, even for one lesson, successfully met discontinued criteria

Notes:

Full series of lessons: Those who received at least 20 weeks of the 30-minute daily lessons or successfully completed lessons before 20 weeks.

Did not receive a full series of lessons: Most often those students who do not begin lessons until late in the year.

Discontinued: Those who successfully meet the rigorous criteria to be discontinued (released) from Reading Recovery during the school year or at the time of year-end testing.

* Does not include Descubriendo la Lectura students. Results for the 3,232 Descubriendo la Lectura students served in 2000–2001 are comparable to Reading Recovery.

Every child served in Reading Recovery, even if for only one day, is counted and reported in data from the National Data Evaluation Center.

These percentages translate into powerful realities. For example, of the 86,009 students scoring at or above grade-level expectations on post-treatment measures of literacy skills, only 113 (0.13%) were placed in special education settings for reading support, and only 23 were assigned to special education settings for writing instruction at the end of first grade. At the same time, only 194 (0.26%) were retained in Grade 1 due to reading performance. In fact, placement in special education or retention in Grade 1 for reading difficulties was found to occur more frequently in the randomly selected, non-treatment group representing average performance than in the group of children successfully served by Reading Recovery.

The children who are not discontinued and are recommended for further evaluation also make progress in Reading Recovery. A review of scores on measures that exhibit a ceiling effect is informative. Measures with a ceiling effect can be treated as criterion measures, that is, performance levels that all first-grade children should reach some time during the first grade. For example, during fall testing for the phonemic awareness measure (Hearing and Recording Sounds in Words), recommended students accurately recorded only 6.6 sounds on average compared to 23.9 on average for the random sample children. In the spring testing, the recommended students recorded 31.8 sounds and the random sample 35.1 sounds. Although this group of recommended students did not meet Reading Recovery's rigorous criteria for being discontinued, the evidence suggests that these children met the criterion level on several measures (NDEC, 2002).

1. Reading Recovery counts every child.

Authors of the Internet letter claim that “studies conducted by researchers associated with Reading Recovery typically exclude 25–40% of the poorest performing students from the data analysis” (paragraph 3). Two possible origins of this argument are hypothesized. First, a 1995 article (Center et al.) asserted that Clay's studies had excluded about 30% of children who were either removed or not discontinued from the program. However, Clay's 1979 data clearly negate this claim: No children were dropped from her analyses. Clay responded to this claim in a published letter in *Reading Research Quarterly* (1997). Yet the Center et al. accusation has been carried forward on an ongoing and inaccurate basis by other researchers.

A second possibility is that the Internet letter authors were referring to studies which have included only discontinued children, those who have successfully completed Reading Recovery lessons. For some research, it is very appropriate to study specific groups of Reading Recovery children to answer identified research questions.

For example, some researchers have studied children who successfully completed lessons in order to determine if children who reach average performance at the end of Grade 1 maintain that average status in subsequent years. This is a very legitimate research question. To answer this question, no researcher would include every child; some children would have received few lessons and comparisons would be inappropriate.

Regardless of the confusion leading to the claim that Reading Recovery excludes poorest-performing students from data analysis, it is important to acknowledge that every child served in Reading Recovery, even if for only one day, is counted and reported in data from the National Data Evaluation Center. All evaluation data are inclusive of all children, regardless of outcome status. The broad accusation made in the Internet letter is misleading at best.

2. Reading Recovery serves first-grade children with the greatest literacy need.

The Internet letter confuses the description of the target population that Reading Recovery serves. No prerequisite skills are required for being served by Reading Recovery. Reading Recovery is a safety net intervention for first graders who are having difficulty with early reading and writing. It is an important component of a school's comprehensive literacy program. Reading Recovery provides instruction for the lowest-achieving children first. Annual Reading Recovery data clearly demonstrate that those served first in the school year have the lowest scores on literacy measures at the beginning of the first-grade year.

All students with the lowest scores for their school enter the program and represent the target population (Clay & Tuck, 1991). Outcomes cannot be predicted reliably by any measure at entry. Any attempt to restrict access to Reading Recovery service based on low entry test scores would deny service to children most in need and many who later successfully complete Reading Recovery lessons.

Reading Recovery has two positive outcomes. First, Reading Recovery students successfully complete lessons (discontinue) when they are judged to have met the criteria to participate in their classroom literacy instruction at an average level and to effectively apply strategies that will support future literacy learning. Second, those students who have not achieved these rigorous criteria after 20 weeks of lessons have still received an intensive period of diagnostic teaching. Recommendation for further evaluation is a positive step because evaluators have much more information about student strengths and weaknesses than was available before Reading Recovery. In addition, students have learned many new

skills and strategies even though they did not meet criteria for successful completion of lessons.

Many Reading Recovery studies disaggregate the data by these categories (discontinued and not discontinued) to provide an indication of the proportion of children returned to average levels of performances versus those identified as needing additional evaluation or alternative support.

3. Reading Recovery methodology follows rigorous standards for evaluation research.

The criticism that Reading Recovery does not follow intent to treat methodology is misplaced (paragraph 3). The intent to treat methodology is used in experiments or trials where research subjects are randomly assigned to one of several treatments. In contrast, Reading Recovery serves children identified by the school as the lowest-achieving first-grade students and reports on the total population served in Reading Recovery. Reading Recovery lessons are not clinical trials with random assignment but rather services delivered to an entire population.

An important aspect of intent to treat, however, is to account for every study subject, and Reading Recovery does meet that criterion. The only students excluded from National Data Evaluation Center reports are those for whom outcome status is missing, and that fact is clearly reported. In the 2000–2001 year, the annual data accounted for 99.93% of children served. Outcome data were missing for only 90 of the 152,241 children served by Reading Recovery and Descubriendo la Lectura (NDEC, 2002).

Data are reported for every child enrolled in Reading Recovery. Students are tested pre- and post-treatment (before and after being served) as well as at year-end. At the end of each child's series of Reading Recovery lessons, a status category is assigned. The five status categories are

- **Discontinued:** The child has successfully completed the program and is able to benefit from classroom instruction.
- **Recommended action after a full program of 20 weeks:** The child was recommended for further evaluation and consideration of other support services.
- **Incomplete program at year-end:** The school year ended before the child had time to complete the program.
- **Moved while being served.**
- **None of the above:** This category is used only in special

circumstances when the child has to be removed from the program for very unusual reasons (such as return to kindergarten).

In 2000–2001, of all 149,009 Reading Recovery students (excluding students in Descubriendo la Lectura) served, regardless of the number of lessons received, 59% discontinued their lessons successfully, 17% were recommended for further evaluation, 16% had incomplete lessons at the end of the school year, 5% moved before lessons could be completed, and 3% were classified as none of the above. Of the children with the full series of Reading Recovery lessons, 78% successfully discontinued their programs. Results for the 3,232 children served in Descubriendo la Lectura were comparable to Reading Recovery results.

C. Bias in the Internet letter is analyzed.

Given all the evidence for the effectiveness of Reading Recovery, how do the authors of the Internet letter conclude that the program doesn't work for the lowest-performing students? First, they limit their view of acceptable evidence to experimental, random control group experiments. This allows them to ignore the extensive data collected annually and publicly reported on the progress of every single student enrolled in the program. To advance their political agenda, the authors of the Internet letter selectively report and distort the limited set of experimental studies they present to their colleagues, politicians, and the public for their support.

The clearest instance of this distortion is in their use of the Elbaum et al. (2000) meta-analysis. They state, “In fact, for the poorest readers, empirical synthesis of ‘in-house’ and independent studies indicates that Reading Recovery is not effective. In Elbaum et al. (2000), the gains for the poorest readers instructed by Reading Recovery were almost zero” (paragraph 3). Why is this a distortion? First, it ignores the major finding of this meta-analysis that the effect of Reading Recovery on student performance was large and significant.

Second, it confuses a post-treatment outcome variable, discontinued versus not discontinued, with a pre-treatment aptitude variable. Students are discontinued from Reading Recovery service when they are judged to have met the criteria to participate in their classroom literacy instruction at an average level and have a set of strategies judged to be sufficient to support future literacy learning. Students who are not discontinued have not achieved these criteria with 20 weeks of lessons and are judged not likely to meet these criteria with an additional week or two of lessons. Many Reading Recovery studies disaggregate the data by these cat-

egories to provide an indication of the proportion of children returned to average levels of performances versus those identified as needing long-term literacy support. Both of these outcomes have positive implications for the efficient use of future educational funds.

The authors of the Internet letter imply that these not-discontinued students were those who entered the program with the lowest scores. This is not the case. All students enter the program with low scores and represent the target population. Again, the Elbaum et al. (2000) study indicates that for the total group the effects of Reading Recovery are large and significant. The outcome category cannot be reliably predicted by any measure at entry. Any attempt to restrict access to Reading Recovery service based on entry tests would deny service to those children most in need of service, many who show substantial benefit from that service.

Elbaum et al. (2000) state, “For Reading Recovery interventions, effects for students identified as discontinued were substantial, whereas effects for students identified as not discontinued were not significantly different from zero” (p. 605). The Internet letter emphasizes the small number of students who did not make progress while it ignores the fact that the majority of students made substantial progress. It also implies that the not-discontinued students were the ones with the lowest entry scores. This is not true. The interpretation of this information in the Internet letter seems biased.

D. Summary

Studies and research reviews in peer-reviewed journals document the effectiveness of Reading Recovery for the lowest-performing first graders. In addition, evidence of subsequent gains is substantial. Moreover, self-esteem studies reveal that Reading Recovery children improve self-efficacy scores as a result of the treatment.

Authors of the Internet letter limit their view of acceptable evidence to experimental, random control group experiments. This allows them to ignore the extensive data collected annually and publicly reported on the progress of every single student enrolled in the program. Then they selectively report and distort the limited set of experimental studies they present.

The National Data Evaluation Center collects, analyzes, and reports data on every child served each year. Results are impressive for these children who initially have the most difficulty with literacy learning. What other program provides such external and internal evidence of success with the lowest-achieving first graders?

Authors of the Internet letter limit their view of acceptable evidence to experimental, random control group experiments. Then they selectively report and distort the limited set of experimental studies they present.