

SECTION II

COST-EFFECTIVENESS IS A COMPLEX CONCEPT IN EDUCATION.

Given the competing demands on education budgets at federal, state, and local levels, it is important to address literacy learning in cost-effective ways. To date, however, no cost-effectiveness analysis comparing alternative early literacy interventions has been conducted (for example, small group instruction compared with one-to-one instruction). This is not surprising since cost-effectiveness analysis is not a widely used evaluation tool in education decision making, although it is often discussed (Hummel-Rossi & Ashdown, 2002; Levin & McEwan, 2001).

The Internet letter states that “Reading Recovery is not cost-effective because the developers require one-to-one interventions by highly trained teachers” (paragraph 4). This assertion could be true only if working with low-achieving children in groups was as effective as working with them one-to-one.

This section

- A. reviews cost-effectiveness issues and the long-term benefits of Reading Recovery,
- B. documents research demonstrating the efficacy of one-to-one interventions for the lowest-performing first-grade students, and
- C. analyzes the quality of research that the Internet letter cites to support claims for superior results of group interventions.

A. Cost-effectiveness cannot be oversimplified.

1. Long-term versus short-term benefits and the population served

To describe any program as expensive or not cost-effective is misleading without providing crucial information such as which students the programs target and what results are sought relative to the performance of other students. Another important

To describe any program as expensive or not cost-effective is misleading without providing crucial information such as which students the programs target and what results are sought relative to the performance of other students.

consideration in cost-effectiveness is the longer-term benefit of the intervention.

In the case of Reading Recovery, the long-term benefits of literacy achievement may significantly outweigh the short-term cost of instruction and teacher preparation. Reading failure is a predictor of academic underachievement. As the achievement gap widens, the potential for future economic, social, and psychological problems increases for this group of children (Pikulski, 1994). The significant number of children—estimated to be 20%—who experience early reading failure are at risk of not reaching acceptable standards throughout their academic careers (Hill & Crevola, 1997). The price of achieving literacy for all students should be weighed against the social ills associated with literacy failure and drop-out rates.

In the case of Reading Recovery, the long-term benefits of literacy achievement may significantly outweigh the short-term cost of instruction and teacher preparation.

The role of prevention (Pianta, 1990) is essential when considering long-term costs associated with literacy failure. Reading Recovery is a secondary prevention that targets an identified group of the first-grade population (lowest-performing) with the highest likelihood of experiencing literacy failure, even within good primary prevention efforts (classroom instruction). By targeting the lowest-achieving first graders, two positive outcomes are possible: a child achieves successful literacy performance or is identified for further assessment and possibly for an appropriate alternative instructional program known as tertiary prevention.

By intervening early, Reading Recovery reduces referrals and placements in special education (NDEC, 2002), limits retention, and has demonstrated lasting effects. Retention and special education referral also have substantial price tags. The local costs of providing Reading Recovery services for 12 to 20 weeks will be substantially less than these costs, particularly when the majority of Reading Recovery children sustain their literacy learning gains.

The Indiana Education Policy Center at Indiana University conducted an independent study of Reading Recovery in the state and reported that the program's impact resulted in reduced grade-retention rates in all schools and larger gains in passing rates on the Comprehensive Test of Basic Skills-5/Terra Nova Form B two years after the intervention in high-poverty schools (Manset, St. John, & Simmons, 2000).

Other outcomes associated with literacy achievement, although not so easily measured or valued, must be considered when evaluating costs. Barnett (1993), in an economic evaluation of human service interventions, called these variables “qualitative residual outcomes” (p. 95). Within the context of Reading

Recovery, a qualitative residual outcome is the investment in teaching skill: teachers can apply their additional skill level in instructional contexts other than Reading Recovery.

Cunningham and Allington (1994) address the complexity of cost-effectiveness analysis.

The criticism most often made of Reading Recovery is that it is too expensive and that it requires too much teacher training. However, getting these results with the hardest-to-teach children leads us to conclude that the teacher training is providing the teacher with extraordinary insight and skills. It does cost money to hire and train Reading Recovery teachers but it also costs money to employ transitional-grade teachers (e.g., pre-first classes), resource room teachers, and remedial teachers, too. It costs money to retain children....When you compare the success rate of Reading Recovery with other programs that keep children for years and never get them reading on grade level, Reading Recovery is a bargain! (p. 255)

2. The difficulty of assessing prevention costs

Costs in any prevention program are difficult to assess because so many factors must be considered:

- regional cost variables affecting salaries, overhead, and more
- level of need for the service and level of coverage provided
- quality of training for teacher leaders and teachers
- efficiency and effectiveness factors in program delivery
- acceptance of program as an integral part of the system
- sustained gains for children resulting in reducing referrals to special education and lower rates of retention at the end of first grade

Most school districts expect that there will be personnel costs and costs for training and materials for every program implemented in the system, particularly programs targeted to the lowest achievers. As Levin and McEwan (2001) point out, decision makers have to realize that some children will cost more to educate. Many districts consider in-service training of teachers to be an ongoing responsibility of a school district and do not consider Reading Recovery training an additional expense.

“When you compare the success rate of Reading Recovery with other programs that keep children for years and never get them reading on grade level, Reading Recovery is a bargain!”

(Cunningham and Allington, 1994)

3. Prevention costs in Reading Recovery

The investment in Reading Recovery reduces the number of children who need ongoing, expensive services. Two school districts have calculated the relative costs of Reading Recovery versus the costs of first-grade retention, Title I remedial instruction, and special education for children classified as learning disabled. These analyses have used district teacher salary figures to calculate both the annual and the cumulative amounts of time that a single child would be likely to spend in each of the programs.

- a. **Lyons and Beaver** (1995) conducted a cost comparison analysis for first-grade retention in Lancaster, Ohio four years after Reading Recovery was implemented system-wide. The study revealed that the first-grade retention rate dropped from 4.3% (76 of 1,772 students) in the three years prior to implementation of Reading Recovery to 2.9% (63 of 2,123 students) four years after systemwide implementation. Using teachers' salaries and students' time in the program, these figures represented a cost savings of \$163,020. In addition, the Lancaster district looked at special education placements. In the three years prior to full implementation of Reading Recovery, 32 students were placed in learning disabilities classrooms at the end of Grade 1 or during the first few months of Grade 2. In the three years after Reading Recovery implementation, 10 children were classified as learning disabled. The cost of educating one learning disabled student at the time was conservatively estimated at \$9,100 across four years of service compared with the per pupil cost of \$1,708 for Reading Recovery. The authors found that considerable savings were realized after the district established Reading Recovery as a prevention program.
- b. **Assad and Condon** (1996) conducted another cost-effectiveness study of Reading Recovery in Fall River, Massachusetts. During a 2-year period (1993–1994 and 1994–1995), the Fall River Reading Recovery program served 186 students at an annual per pupil cost of \$2,363. Added to this was the cost of additional interventions for several referred or retained children, for a total implementation cost of \$483,271. Using the data collected on retention, special education, and Title I placement in years prior to Reading Recovery implementation, district administrators estimated that without Reading Recovery, 50% of the Reading Recovery students would have been referred to special education and 50% would have been referred for Title I services. Administrators also estimated that approximately 5.7% would have been retained. Using these figures, district administrators estimated a 5-year cost of

\$1,746,145 if Reading Recovery had not been implemented in the district, for a net savings of \$1,262,874. This dollar amount, however, does not translate directly into a reduction of school department spending. It is an estimate of the resources that because of Reading Recovery will not be needed for teaching basic literacy skills, thus allowing funds to be shifted to meet other important needs.

B. One-to-one interventions are essential for accelerated learning of the lowest-performing first graders.

1. Evidence supports one-to-one intervention for the lowest-achieving first graders.

One-to-one intervention is more reliable than group programs. Solid scientific evidence supports the effectiveness of Reading Recovery's one-to-one tutoring model versus small group instruction for the lowest-performing first graders (Dorn & Allen, 1995; Pinnell et al., 1994). The individualized tutoring enables the highly trained Reading Recovery teacher to tailor each lesson to the unique needs of each struggling student. This individual tutoring, in contrast to rigidly scripted programs for all children, is efficient because the teacher does not waste time on what the child already knows. The Reading Recovery framework is qualitatively different for every child because the teacher makes decisions based on individual needs within each lesson component. The teacher is always pushing the boundaries of the learning of the particular child. This explicit and intensive instruction would be weakened if teacher time was divided among several other children.

Evidence that small group instruction is just as effective with this challenging group of young learners is seriously suspect. For example, the Elbaum meta-analysis cites only two studies that compared a one-to-one intervention with a small group intervention (Acalin, 1995; Evans, 1996). Neither study justified the claim made for small group instruction in the Internet letter (see comments on both studies on pages 39 and 40 of this report).

The assertion of superior cost-effectiveness of small group instruction goes against years of research documenting failure of traditional small group remedial instruction and ability grouping to close the gap for children from poor and minority backgrounds (Allington, 2001; Allington & Cunningham, 2002; Shepard, 1991). One-to-one tutoring of primary-grade students has been demonstrated to be a successful strategy (Wasik & Slavin, 1993).

Arguments related to one-to-one instruction versus group instruction must acknowledge the difficulties of comparing effects of

The assertion of superior cost-effectiveness of small group instruction goes against years of research documenting failure of traditional small group remedial instruction and ability grouping to close the gap for children from poor and minority backgrounds.

these two delivery systems. For example, many factors could influence the outcomes in both one-to-one and small group settings, including age of subjects, initial performance of subjects, measures used, criteria for success, size and design of the study, duration of the intervention, curriculum content, quality of training and pedagogy, length of intervention, subsequent performance, and more.

Hurry's (2000) review of intervention research states, "The available evidence on the effectiveness of one-to-one tuition [teaching] is very positive, but the curriculum content and pedagogy are also important" (p. 20). Pinnell and her colleagues (1994) make a similar argument: One-to-one setting, a lesson framework with intensive literacy experiences, and long-term teacher training are all necessary but not sufficient to explain the success of Reading Recovery.

We know that not all one-to-one interventions and not all small group programs yield positive outcomes for children; therefore, consideration of complex factors contributing to success in both delivery systems merits attention. It is not simply a question of one-to-one versus small group instruction.

2. The MacArthur Foundation-funded study (Pinnell et al., 1994) supports one-to-one intervention.

This well-designed, large scale experimental field study (40 schools) was designed in response to challenges about the delivery system of Reading Recovery. Reading Recovery was systematically compared with (1) another one-to-one intervention, (2) a one-to-one intervention with teachers who had limited training in Reading Recovery, and (3) group instruction based on Reading Recovery principles with trained Reading Recovery teachers.

The lowest children (N=324) in the 40 schools were randomly assigned within schools to either one of the four treatments or a control group. Researchers at the University of Chicago independently analyzed the data. In addition, a renowned national panel of researchers not involved in Reading Recovery provided oversight for analyzing the results.

Results were definitive. Reading Recovery subjects performed significantly better than any other treatment and comparison groups on all measures. Essential differences were related to individual instruction, the lesson framework (combination of techniques), and teacher training. (See pages 19–20 for further description of results.)

3. An Arkansas study (Dorn & Allen, 1995) supports one-to-one intervention.

Dorn and Allen reported the simultaneous implementation of Reading Recovery and a specially designed small group model. Extensive staff development was provided to Reading Recovery teachers who taught Reading Recovery (30-minute sessions) for part of the day and small groups (45-minute sessions) for part of the day.

The lowest children were served first in Reading Recovery; others were placed in groups of five. When a child exited Reading Recovery, the lowest child in a small group or the lowest child from a classroom was placed in Reading Recovery. Priority was given to offering the individual tutoring for children who needed it most.

Data for 231 children were analyzed: 95 received Reading Recovery tutoring only, 93 received group services only, and 43 received a combination of group service and Reading Recovery. Of the children receiving Reading Recovery only, 76% were successfully discontinued from service. These were initially the lowest children in the study. Of the children receiving only group service, 30% reached successful levels of reading achievement. These children were initially higher than the Reading Recovery group in reading performance. Dorn and Allen concluded that Reading Recovery was the most effective program for the lowest children who must have individually tailored lessons.

C. Research cited by the Internet letter is flawed.

To support its recommendation for change in Reading Recovery group size, the Internet letter cites a meta-analysis (Elbaum et al., 2000) and two unpublished doctoral dissertations (Evans, 1996; Iversen, 1997). Although the previous section reviews some aspects of the Elbaum meta-analysis, this section focuses on problematic findings related to group size, group composition, and fidelity of treatments included in the meta-analysis.

1. Elbaum meta-analysis

Elbaum, Vaughn, and Moody (2000) attempted a synthesis of 31 different studies. The goal was to explore relative effects of various features of intervention programs. These features included, among others, small group versus individual instruction and Reading Recovery versus other types of interventions. Of the 31 studies

Of the children receiving Reading Recovery only, 76% were successfully discontinued from service. These were initially the lowest children in the study. Of the children receiving only group service, 30% reached successful levels of reading achievement.

(Results of 1995 study by Dorn and Allen)

included in the analysis, 10 purported to include Reading Recovery, but at least two of them (McCarthy, Newby, & Recht, 1995; Iversen & Tunmer, 1993) did not deliver services meeting national Reading Recovery standards.

A significant flaw exists in the Elbaum et al. meta-analysis design. In order for meta-analysis results to be fair and effective, studies must include substantially similar constructs (Wortman, 1992). According to the writers, 31 studies included “adult-delivered one-to-one instructional interventions in reading for elementary school children identified as being at risk for reading failure” (p. 606). Fourteen of these studies examined tutoring of children beyond first grade. Reading Recovery includes only first-grade children. Within the 29 treatment-control studies in the meta-analysis database, Elbaum and colleagues identified 14 different constructs with a total of 54 different variations. For example,

- At-risk was variously defined as lowest 20th–30th percentile or learning disabled.
- Instruction was delivered by teachers, college students, or volunteers.
- Five different foci of instruction were defined across the various studies.
- Multiple and complex outcome measures were used in the various studies.
- One-to-one instruction was variously compared to classroom instruction and other one-to-one comparison groups.
- Except for the Reading Recovery feature, the authors did not separate the studies of first-grade interventions from those that looked at older children.

Another problem area in the Elbaum et al. meta-analysis is the treatment of the study by Pinnell et al. (1994). This large-scale field study compared three one-to-one treatments (Reading Recovery, Reading Recovery-like tutoring with a partially trained teacher, and a skills tutoring model) to one group treatment (a small group taught by a Reading Recovery teacher). Elbaum inappropriately averaged the effects of the three experimental one-to-one treatments. These three one-to-one treatments were very different approaches and had different outcomes; therefore, averaging was inappropriate.

As evidence that Reading Recovery as a one-to-one tutorial had no advantage over small groups, Elbaum et al. cite an unpublished doctoral dissertation by Evans (1996) and an unpublished master’s

thesis by Acalin (1995), although the reference list in Elbaum et al. mistakenly describes Acalin's work as an unpublished doctoral dissertation. Even a casual reading of these two studies reveals that they do not qualify as investigations of Reading Recovery interventions. In fact, a closer reading of these two studies identifies serious problems.

2. Evans unpublished doctoral dissertation

Both the Internet letter and the Elbaum et al. meta-analysis use Evans' unpublished doctoral dissertation (1996) as evidence that Reading Recovery provides no advantage over small group instruction. Yet the Evans study is an unpublished qualitative doctoral dissertation designed to explore literacy acquisition and peer interaction based on eight case study descriptions. The entire study consisted of eight children: four randomly assigned to Reading Recovery and four assigned to a small group intervention. The researcher was also the teacher of the small group intervention, and the Reading Recovery teacher was in the first months of her Reading Recovery training year. The researcher-small group leader "had taught young children in public elementary schools for seven years" (p. 29), while the Reading Recovery teacher's experience "had primarily been in third through fifth grades" (p. 29–30) before beginning Reading Recovery training. The study was conducted between August and December, at the beginning of the Reading Recovery teacher's training.

Evans stated that the eight children had similar entry scores but did not provide data for selection or matching subjects. The summary table of average pre- and post-test scores (p. 129) listed lower pre-test scores and larger standard deviations for the Reading Recovery treatment group across all six subtests (with the exception of an equivalent pre-test score for text reading level, at 1.25). Yet Elbaum and colleagues reported a mean effect size for Evans' study despite the lack of evidence of equivalence at pre-test.

3. Acalin unpublished master's thesis

The stated purpose of Acalin's (1995) study was to compare the effectiveness of Reading Recovery to Project Read. Project Read is a remedial reading program "originally designed to be used in kindergarten through eighth grade" (p. 20). The results of this study are suspect because of the training provided. The specialists who provided the Project Read instruction "had received the full training in this program and were following the manuals, lesson by lesson, with minimal program adaptations" (p. 35). Yet the Reading Recovery treatment was not delivered by Reading Recovery teachers, but by special education teachers who had not

A body of research supports one-to-one tutoring and indicates that it may be essential for children who are at high risk.

participated in the required Reading Recovery training. Descriptions of the instructional methods used for Reading Recovery in this study indicate wide discrepancies from published Reading Recovery procedures.

Further, although Reading Recovery is a first-grade intervention, Acalin provided instruction to 66 subjects ranging from kindergarten through fourth grade and included students who were already placed in a resource specialist program and identified as learning disabled. Only 8 of the 66 children were in the first grade, and four of these children were assigned to each approach.

Even casual consumers of scientific research would wonder why the Evans and Acalin studies were considered to have met criteria for inclusion in a meta-analysis that purports to follow, in the words of Elbaum et al., “best practices for research synthesis” (2000, p. 606).

4. Iversen study

The Internet letter cites Iversen’s (1997) study to claim no advantage of one-to-one instruction over small group instruction. In this study, 75 first graders were grouped in triads (without random assignment). Each trio was taught by the same teacher, one individually and the others in a group of two, for a maximum of 60 lessons. Iversen concluded there was no advantage of individual over small group instruction. Yet the Reading Recovery intervention in this study was far from standard Reading Recovery. There were major differences in the training model, the procedures related to selection of children, teaching procedures, and issues of implementation and evaluation practices. In addition, there are questions about design and methodology.

Because Iversen’s study used a modified version of Reading Recovery, it seems the Internet letter authors are mixing two very different researchable issues: (1) standard Reading Recovery versus a small group intervention and (2) any one-to-one intervention versus a small group intervention. To examine Reading Recovery versus small group instruction, the standard program deserves to be studied. Therefore, the design of this study does not allow any comparisons of group instruction to standard Reading Recovery.

D. Summary

Cost-effectiveness is a complex concept that cannot be oversimplified. It is misleading to describe any program as expensive or not cost-effective without first identifying crucial information such as

which students the program targets and what results are sought relative to the performance of other students. Both long- and short-term benefits must be considered.

In the Internet letter, concerns about Reading Recovery costs focus primarily on one-to-one instruction and the highly trained teachers Reading Recovery requires versus group instruction. Solid evidence (Dorn & Allen, 1995; Pinnell et al., 1994) supports the effectiveness of Reading Recovery versus small groups for lowest-performing first graders. Evidence that small group instruction is as effective with this group of learners is seriously suspect.

Studies cited by Elbaum et al. provide virtually no evidence to support a change from one-to-one to small group instruction for the lowest-achieving first graders. The suggestion for Reading Recovery to change from one-to-one instruction is especially weak because there is documented evidence of effectiveness with hundreds of thousands of children. Certainly, researchers within and outside Reading Recovery should continue to study all possibilities, but the idea of change in group size needs a much stronger research base. A body of research supports one-to-one tutoring and indicates that it may be essential for children who are at high risk (Bloom, 1984; Juel, 1991; Wasik & Slavin, 1993). The systematic nature of Reading Recovery instruction is based on a teacher's detailed assessment and analysis of a child's knowledge base and skills. The teaching is highly efficient because the teacher has this precise inventory of skills and strategies and is able to teach exactly what the child needs to know next.

To use their own words, Elbaum and colleagues conclude, "In sum, the findings of this meta-analysis support the argument that well-designed, reliably implemented, one-to-one interventions can make a significant contribution to improved reading outcomes for many students whose poor reading skills place them at risk for academic failure" (2000, p. 617).