

Reading Recovery in the United States: More than a Decade of Data

Carol A. Lyons

The Ohio State University

Abstract

There is considerable information available to evaluate Reading Recovery's impact on children's literacy development and the professional development of teachers. The purpose of this article is to review the thirteen years of replication data that support Reading Recovery's effectiveness, as well as to address the questions most often raised by critics regarding (a) the length of the teacher training program, (b) the cost of implementation, and (c) the long-term effects of the program for children. Rationales are explicated for leaders of the program requiring that certified teachers enroll in a year of academic coursework and participate in continued professional development, teach the lowest achieving children one-on-one, and collect and report data on a daily basis to document the effectiveness of the program.

Reading Recovery (RR), an early intervention literacy program, has been operating in the United States for more than a decade. During that time, nearly a half million children have received instruction and 15,000 teachers have participated in training. Because of the extent of the development of this program, there is considerable information available to evaluate its impact on children's literacy development and the professional development of teachers.

The purpose of this article is to review the extensive replication data that support RR's effectiveness, as well as to address the questions most often raised by critics regarding (a) the length of the teacher training program, (b) the cost of

implementation, and (c) the long-term effects of the program for children. After a brief history of the development of RR, each of these areas will be discussed separately, and there will be a general call for programs to substantiate their effectiveness in the quest toward literacy for all children.

A Brief History of Reading Recovery

In September 1984, Professor Marie M. Clay, a New Zealand researcher and educator who originally designed the program, and Barbara Watson, current National Director in New Zealand, introduced RR to faculty at The Ohio State University and sixteen teachers in the

An International Journal of Early Reading and Writing™
An Official Publication of the Reading Recovery Council of North America

Columbus Public Schools. This early intervention program provided intensive, individual help to the lowest achieving first grade students in six Columbus, Ohio schools. End-of-year data revealed that during the initial year of implementation, when all of the educators were learning the program, 67% of the lowest achieving children developed effective strategies for reading and writing and reached average classroom levels after 12-20 weeks of one-to-one instruction.

In July 1985, the successful results of the pilot study led the Ohio General Assembly to provide funding to establish teacher training sites in Ohio and to begin implementing the program throughout the state. By the start of the 1987 school year, RR was operating in 182 school districts throughout Ohio. When the number of low-progress first grade children who were reaching average reading levels increased from 73% in 1986 to 79% in 1987, long-term benefits of the RR program became a possibility. The Ohio General Assembly and the Ohio Department of Education have continued to fund the training and ongoing professional development of RR teachers and teacher leaders for 12 years.

In 1987, the U.S. Department of Education's National Diffusion Network (NDN) recognized RR as an exemplary research-based program and provided funding to make the program available to school districts in other states. Four educators from outside of Ohio enrolled in the year-long RR teacher leader course at The Ohio State University during the 1987-1988 academic year. These educators returned to their home sites the following year to begin training teachers to deliver the program to children. In 1996-97, the United States RR program was

operating in 48 states, the District of Columbia, and some U.S. Defense Department Schools overseas. As reported in Table 1, the RR network by 1996-1997 includes 42 university trainers, 667 teacher leaders, 15,483 teachers, 3,241 school districts, and 9,815 schools.

Reading Recovery Evaluation

Data:

Replication Methodology

Replication of results represents a vital component of research and an important concept in the history and theory of research design (Campbell & Stanley, 1963; Kratochwill, 1978). Intervention research in such fields as medicine, social work, psychology, and education seemingly requires replication of results to an even greater degree (Hersen & Barlow, 1976; Neuman & McCormick, 1995). There are two main approaches to replication: (a) Systematic replication, which involves different investigators conducting the same study on different subjects with the same problem at a different location and at a later time, and (b) simultaneous replication, which is similar to the former, but being conducted at the same time (Gay, 1987).

While replication at a later time (i.e., systematic) is the standard approach, simultaneous replication, a research methodology designed by Frymier, Barber, Gansneder, and Robertson (1989), has been used effectively to assess students' academic achievements in widely separated geographical settings. For example, simultaneous replication was successfully tested as a methodology in the Phi Delta Kappa Study of Students at Risk (Frymier et al., 1989) by subjecting to common analytic procedures data that were collected in common ways with

common instruments in 87 separate sites. The study enabled researchers from urban, rural, and suburban school districts to verify results by focusing on local analyses of data collected at the school level while still contributing to a large-scale study. To determine program effectiveness, RR has employed both types of replication methodology.

Data evaluating the original design of the program are monitored annually in New Zealand by the Ministry of Education (Kerslake, 1996). Since 1985 in the United States, the effect of the RR program has been replicated hundreds of thousands of times in thousands of schools with hundreds of thousands of individual subjects. In that time, approximately 15,000 RR teachers working individually with more than 435,000 low achieving first grade children from different cultures in urban, suburban, and rural school districts have documented similar results. That is, RR teachers, using the RR teaching procedures they learned through standardized professional train-

ing, have helped the lowest achieving first grade students reach average band reading levels after 12 to 20 weeks of individually designed and individually delivered instruction (Lyons, 1997). Essentially, children who were initially labeled "slow" learners were shown to be performing at average levels in reading, and some reports have indicated profits in other subject areas, as well (Lyons & Beaver, 1995). Furthermore, RR teachers in other countries such as Australia, Canada, The United Kingdom, New Zealand, and the U.S. Defense Department Schools overseas have produced notably similar results. According to Kratochwill (1978), repeatedly producing the same effect with different students in different settings increases confidence in a treatment or intervention, thereby providing substantial evidence of the effectiveness of RR tutoring.

As reported in Table 2, from 1985-1997, the RR program served a total of 436,249 children. Of that group, 313,848 had sufficient time to experience a com-

Table 1 U.S. University Trainers, Teacher Leaders, Teachers, School Districts, and Schools Participating in Reading Recovery from 1984-1997

Year	University Trainers	Teacher Leaders	Teachers	School Districts	Schools
1984-85	0	0	16	1	6
1985-86	1	3	58	23	35
1986-87	3	2	280	108	255
1987-88	3	45	531	143	227
1988-89	6	43	732	265	623
1989-90	11	54	1,163	332	892
1990-91	13	80	1,850	508	1,406
1991-92	19	155	3,164	798	2,336
1992-93	24	259	5,343	1,246	3,731
1993-94	33	388	8,182	1,905	5,523
1994-95	39	510	12,084	2,543	7,784
1995-96	39	625	14,153	2,939	9,062
1996-97	42	667	15,843	3,241	9,815

Note. Data from the National Evaluation Data Center, The Ohio State University

plete program (defined as 60 lessons) and 81% reached criteria for successful release from the program; that is, they were performing within the average band reading group of their classroom. Such numbers represent extensive replication documentation, a hallmark of research reliability. Data documenting the impact of RR on student achievement are reported each year in local, state, and national evaluation reports. In addition, several other reports produced by RR professionals and others provide data to document three claims related to the effectiveness of the RR program.

Claim 1: Within 12-20 weeks of daily, one-to-one instruction, the majority of the lowest achieving first grade students can be placed in an

average reading group in their respective first grade classrooms. Since 1984, data for every child served in the U.S. have been reported to the Reading Recovery National Data Evaluation Center at The Ohio State University and forwarded to the United States Department of Education. If we consider all students served, even for one day, 60% met the stringent exit criteria for success.

There is no checklist of specific criteria to determine that a child is ready for discontinuing because the goal of the program is to place the child in a classroom reading group in which he or she is predicted to make progress without further individual instruction. The level of performance will differ from child to child

Table 2 U.S. Reading Recovery Children Served, Program Children and Percentage of Children Discontinued from 1984-1997

Year	Served**	Program***	Discontinued****	%
1984-1985*	110	55	37	67%
1985-1986	230	136	99	73%
1986-1987	2,048	1,336	1,059	79%
1987-1988	3,649	2,648	2,269	86%
1988-1989	4,772	3,609	2,994	83%
1989-1990	7,778	5,840	4,888	84%
1990-1991	12,605	9,283	8,126	88%
1991-1992	21,821	16,026	3,499	84%
1992-1993	36,443	26,582	22,109	83%
1993-1994	56,077	40,493	33,243	82%
1994-1995	81,220	57,712	46,637	81%
1995-1996	99,617	71,193	59,266	83%
1996-1997	108,876	78,935	65,551	83%
Totals	436,249	313,848	259,777	81%

Note. Data from the National Evaluation Data Center, The Ohio State University

*Pilot year: RR teachers were in training.

**Served: Program children and children who entered Reading Recovery but did not receive a minimum of 60 lessons because they moved, were absent for extended periods of time, or the school year ended prior to completion of lessons. Column 1 is inclusive of the subcategory Program Children, column 2.

*** Program: RR children who received a minimum of 60 lessons or were discontinued prior to receiving 60 lessons.

****Discontinued: RR children who were released from the RR program reading within average band reading levels of the class.

and from school to school (Clay, 1993). If the child is to continue to make progress, however, RR teachers must consider whether that child has acquired a system of strategies that helps him or her learn from further attempts to read and write. This system of strategies includes the ability to (a) control left to right directional movement, (b) match spoken to written words, (c) notice and correct errors when reading and writing, (d) notice discrepancies in responses by cross-checking one source of information (e.g., visual) with a different source of information (e.g., meaning or structural cues), (e) use many sources of information, and (f) detect and self-correct errors (Clay, 1993). If RR students acquire these strategies, assuming these strategies are the ones beginning readers must acquire, they should continue to make average progress in reading in the years after they complete RR.

Claim 2: Reading Recovery is more effective when compared to traditional one-to-one and small group remedial programs targeting low-achieving first grade students.

Researchers Wasik and Slavin (1993) compared RR to four other one-to-one tutoring models that have been used to improve the reading skills of first graders who were at risk of failure: Success for All, Prevention of Learning Disabilities, The Wallach Tutoring Program and Programmed Tutorial Reading. Sixteen studies evaluating the effect of these models on student achievement revealed substantial positive effects of one-to-one tutoring in comparison to small group instruction. Wasik et al., (1993) reported "follow-up studies found that effects of tutoring were generally lasting and the results were more positive when reading

instruction was based on a more comprehensive model of reading and when certified teachers (rather than paraprofessionals) were the tutors" (p. 178). The researchers also reported that RR is the only program that has documented long-term success without additional intervention and the only program that has assessed the quality of implementation across tutoring sessions and the effect this has on outcome data. Wasik and Slavin (1993) concluded that when compared to other one-to-one interventions, RR is at least as effective as the others, but one well-wrought study found it is more effective.

In 1988, the John D. and Catherine T. MacArthur Foundation funded an experimental study (Pinnell, Lyons, DeFord, Bryk, & Seltzer, 1994) designed to compare the effect of RR to two alternative one-to-one treatments and one small group treatment, with Title 1 programs as control groups. The results of the study indicated that RR was the only group for which the mean treatment effect was significant on four measures: Dictation 2 (Clay, 1993); text reading (Scott Foresman, 1979); and two standardized tests, the Gates-MacGinitie Reading Test (1989) and the Woodcock Reading Mastery Test (1990). While the effectiveness of RR is not challenged, we need to know more about qualitative differences that exist between the teacher-student interaction during RR lessons and how to maximize Reading Recovery's effectiveness while minimizing the cost.

Claim 3: Reading Recovery greatly reduces the number of children targeted for placement in learning disability (LD) classrooms. Several studies have demonstrated that once RR is introduced into a school system, there is a sharp

decline in the number of first grade students referred for learning disability screening and placement. For example, Lyons and Beaver (1995) reported that in the State of Ohio the number of RR program students referred for learning disability screening decreased from 1.26% to just 0.51% over a five year period. Furthermore, a national study demonstrated that the number of first grade students targeted to receive LD services was cut in half two years after RR was implemented. Specifically, prior to RR implementation, 59 (2.3%) of the 2,569 first grade students in the ten districts were referred for LD services. One year after RR intervention, 53 (2%) of 2,602 first grade students were referred for LD services and only 34 (1.3%) of 2,572 students were referred for LD services two years after the program was implemented (Lyons, 1994). Data reported to the United States Office of Education (Lyons, 1997) for the 1995-1996 academic year, indicated that only .02% of the 71,193 RR children who received full programs were referred for LD assessment.

Research suggesting that RR has the potential to reduce the escalating number of students retained and referred for learning disability testing and placement was cited in a report to the International Reading Association (IRA). The IRA report, *Learning Disabilities - A Barrier to Literacy Instruction* (1995), stated that "RR effectively teaches children to read . . . Not only does it reduce the number of children who are labeled with learning disabilities, but it also significantly reduces the number of children who are retained in remedial reading programs" (p. 45). Furthermore, the program enables educators to separate first grade children who may be low achieving from those

with more severe learning problems (Lyons, 1994).

Essentially, there is no question that RR works. Consider, for example, that Hiebert (1994) noted " . . . a high percentage of Reading Recovery tutees can orally read a first-grade text at the end of Grade 1" (p. 21). Shanahan and Barr (1995) concluded, "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" (p. 989). Most strongly, in *Classrooms That Work*, Cunningham and Allington said, "No other remedial program has ever come close to achieving the results demonstrated by Reading Recovery" (1996, p. 254). What is most often questioned, however, is the need for year-long training and continued professional development of teachers, the costs associated with the program, and evidence of the long-term effects of the early intervention using standardized measures. In the following sections I will address these three issues.

Year-Long Training and Continued Professional Development

Unlike many other programs for low progress students, RR is not based on one procedure or a set of materials to use for instruction. Rather, it is dependent on the customized instruction designed by a specially trained teacher who has developed a systematic knowledge and understanding of possible progressions in acquiring a reading and writing process. The teacher assists the reader in acquiring the strategies employed by successful readers.

own reading and writing. During each lesson, the teacher carefully observes the child acting on a variety of texts and systematically records these observations to form the basis for the next lesson. Learning such a complex role takes time, commitment, much energy, and a rigorous training program.

Teachers in training continue working full-time in their school as they receive instruction in RR techniques. The most common arrangement during the training year and subsequent years is for teachers to spend one half day teaching RR students and the other half day performing other assigned duties (e.g., kindergarten, first grade, Title 1 teacher). Reading Recovery teachers working with four or five students for one half day will teach a total of 7 to 10 children, on average, every year. This represents 14-20 children per 1.0 FTE annually.

Following the training year, teachers meet several times annually with their teacher leader for continued professional development. In these sessions, teachers sharpen their observation skills and learn how to use these observations to design efficient lessons that will most effectively accelerate students' progress. Teachers are also given the opportunity to attend the annual national RR conference and a number of regional professional development institutes to further their theoretical and practical understandings of the reading and writing processes.

Leading authorities in school reform have recognized the quality of the RR training. In a discussion of the use of standards and assessments to support student learning, Darling-Hammond and Falk (1997) singled out RR as effective in helping students gain skills that make them successful and confident readers,

In order to implement the RR program, qualified teachers enroll in a year-long course taught by a certified teacher leader at a training site in or near their school district. Through close observation of teacher and student interactions, guided by a skilled teacher leader, RR teachers learn to use observation techniques to determine where the student's literacy processing is breaking down and why. Extensive use is made of a one-way glass for demonstration and observation. By observing each other working with children behind the one-way glass, teachers become sensitive observers of children's reading and writing behaviors and develop skill in making moment-to-moment teaching decisions that help children use what they know to generate further understandings.

The RR teacher is responsible for teaching children who, despite one year of kindergarten, remain at the lowest achieving level of the first grade class. In order to accomplish this feat, teachers must customize every lesson to meet the idiosyncratic needs of the child by selecting from a wide range of books and helping individuals use their writing to assist in reading. Teachers also perform and record their own assessments of a student's progress in reading. During reading and writing tasks, teachers must select from an array of special techniques those that will help children develop effective problem-solving strategies that independent readers use. Students are taught how to predict, confirm, and understand what they read using all sources of information (e.g., visual, semantic, etc.) As they write, they develop strategies for hearing and recording sounds in words, composing messages, and for monitoring and checking their

including students whose first language is not English and many who have been identified for special education. Allington and Cunningham (1996) noted that "Planned professional development of this intensity is rarely encountered in school improvement efforts" (p.32). The knowledge and skill of the trained teacher is the critical element to RR; the element that distinguishes RR from other programs designed for low progress children; the element that may very well be the deciding factor that allows for the program's success.

After conducting a three-year study of RR, Kenneth Wilson (Wilson & Daviss, 1994), a Nobel Prize winning physicist and educational reformer, concluded that in three ways the program can encourage the process of educational redesign:

First, it proves that a well-designed educational program can be replicated among teachers and schools across a wide array of locations and cultures and still yield uniformly superior results. Second, it indicates that an investment of money and effort in educational design can earn dramatic rewards — if it's made in a properly researched and designed program that offers thorough teacher training and support. Third, it shows that when educators find a program that meets these two criteria and proves that it can earn a good return, schools are willing to make its adoption a budget priority. Reading Recovery is the best evidence yet of the direct link between good design and educational excellence. (p. 76)

American society prides itself on the advancement of technical skills in medicine. Physicians are expected to engage in life-long learning through continuous professional development; some physicians (e.g., surgeons) require more

advanced, long-term training than others. You would not want a surgeon performing a heart transplant on a loved one using the same techniques he or she learned 20 years ago in medical school. You would expect the surgeon to use more effective, proven procedures that he or she learned in advanced surgical training institutes. Advanced life-long professional development for teachers is rare. Some teachers continue to use the same teaching methods they learned in undergraduate teacher education programs.

Reading Recovery teachers should have more specialized continuous professional development because they are required to work with the most difficult to teach students. In order to do so successfully, teachers must learn specialized skills which require specialized training. The RR initial and ongoing professional development program for teachers breaks away from the expected norm. In doing so, there are long term gains for school districts, administrators, teachers, students, and parents.

The best investment this nation can make is in massive ongoing professional development for teachers. Renewing, re-educating, extending, and enhancing the professional expertise of the teachers who carry out the daily work of educating children is critical to school reform. In her presidential address at the American Education Research Association annual meeting, Linda Darling-Hammond (1996) stated that "recent research illustrates that money makes a difference in the quality of education, especially as it is used to pay for more expert teachers, whose levels of participation and skill prove to be the single most important determinant of student achievement (Armour-Thomas et al., 1989; Ferguson, 1991). Furthermore,

students' right to learn is directly tied to their teachers' opportunities to learn what they need to know to teach well" (p. 6).

The Cost of Reading Recovery

In *No Quick Fix: Rethinking Literacy Programs in America's Elementary Schools*, Allington & Walmsley (1995) concluded that "... the more expensive RR program provides the best evidence of long-term success for the largest population of at-risk students served" (p. 262). However, some researchers (Hiebert, 1994; Shanahan & Barr, 1995) report RR costs too much. Others (Dyer & Binkney 1995; Lyons & Beaver, 1995, Moriarty, 1997; Pinnell, Lyons & Jones, 1996) argue that it costs much less than retention and long-term placement in learning disability, special education, or remedial reading resource rooms. Furthermore, the initial start-up cost of the program (i.e., teacher leader training, installation of a one-way mirror, tuition, books and materials, and the initial training of RR professionals) is a one-time expense.

Can the cost for RR be justified?

Those who agree it can be justified weigh the cost of a 30-minute RR lesson for 12 to 20 weeks against the cost of 45 minutes of daily remedial reading groups for more than 1 year. They weigh the cost of a 30-minute RR lesson for 12 to 20 weeks against 5 hours of daily learning disability (LD) classes for 4 to 5 years. They weigh the cost of a 30-minute RR lesson for 12 to 20 weeks against a year of repeating first grade. Educators can expect to spend about 50 percent more to educate a low achieving child (Levin, 1989). The Massachusetts State Legislature reached a similar conclusion after conducting a study of five years of special education placements in the state.

The study revealed that between FY 1990 and FY 1995, total enrollment in special education increased by 8.3% statewide (MA Superintendents Association Task Force, 1997).

Furthermore, an examination of the relative cost of the increased enrollment in regular and special education during this period revealed that expenditures per full time equivalent (FTE) enrollments in special education increased by almost \$4,000 from FY 1990 to FY 1995, while they increased by only \$305 in regular education. The impact of these increases statewide has been dramatic, resulting in an additional expenditure of \$61 million on special education in FY 1995 alone. The report also revealed that expenditures for special education increased at a greater rate than expenditures for regular education in 71% of the Massachusetts school districts with only 3% of the districts reporting a decline in special education expenditures between FY 1990 and FY 1995.

A cost-effectiveness study of special education referrals, Title 1 placement, and retention was conducted in Fall River, Massachusetts (Assad & Condon, 1996). The report revealed that over a two-year period (1993-94, 1994-95) the Fall River RR project served 186 students at an actual cost of \$2,362 per pupil. Based on school history, it was estimated that without the RR program, 45%-50% of the students would have been referred for Special Education; 45%-50% would have been referred for Title 1 services and approximately 5.7% of the students would have been retained. Total cost for special education services in this school district is \$17,830 per pupil; total cost of Title 1 services per pupil is \$4,860; total per pupil cost for retention is \$3,843.

Using this information, Fall River administrators determined a projected five year cost of \$1,746,145 if RR had not been operating in the district. The RR intervention for this same five year period would cost \$483,271, creating a net savings of \$1,262,874 that could be reallocated for a variety of other services needed for students within the district.

Similar cost savings were calculated in Medford, Massachusetts. Data collected in this urban school district over a five-year period revealed that five of the 175 first graders who were successfully released from RR, representing fewer than 3%, have been referred to Special Education. Prior to RR implementation, it was estimated that the vast majority of these students would have been targeted for special education (Moriarty, 1996).

After conducting a seven month investigation, the Massachusetts legislative team concluded that RR research shows a high degree of success in teaching low progress children how to read and write; defers children from special education, reduces the number of children retained, and is cost effective because for every \$3 invested in RR, a school system saves \$5. In 1997, the Massachusetts legislature allocated \$500,000 for early intervention and identified RR as a research-based program that would qualify for funding (Moriarty, 1997).

When examined as a whole, the net costs of RR are justified by the value of all that is saved. First of all, the program is producing effective results time and time again, as evidenced by replication data. Second, the program cuts the cost of retaining a child or placing a child in a learning disability resource room for up to 4 and 5 years. Third, the program cuts

the cost of long-term help in remedial reading resource rooms. Fourth, in preventing more serious problems from occurring, the program cuts the cost of ongoing expensive psychological assessment and treatment. The school district saves money in the long run. To all of these monetary savings, however, must be added the incalculable value of what the program does for the thousands of boys and girls who are spared from a lifetime of feeling inadequate because they cannot read and write well enough to keep up with peers and benefit fully from classroom experiences.

Evidence of the Long-Term Effects of Reading Recovery Using Standardized Measures

Even since the early years of RR in the U.S., there has been interest in determining whether children who are successfully released from the RR program continue to make average band progress in reading and writing. In 1988, the Ohio Department of Education commissioned an outside evaluation team to evaluate long-term effects of the RR program (Anderson, 1988). Over a three-year period (1984-1987), the evaluation team examined the effect of RR on the lowest achieving first grade Ohio children's reading progress. The evaluation team was comprised of nationally known experts in literacy and chaired by Dr. Richard Anderson, at that time the Director of the Center for the Study of Reading at the University of Illinois. The report revealed that 81.8% of RR children who received a full program made accelerated progress and performed within the average band range for their classes. Furthermore, Anderson (1988) found that children "retain their gains and con-

tinue to make progress at least 2 years after the intervention" (p.42).

The long-term effects of any intervention are difficult to measure because there are many intervening variables which can influence children's progress (e.g., quality of subsequent classroom instruction, promotion and disciplinary policies, student's health, mobility, and individual life circumstances). As a matter of fact, few implementers or proponents of any intervention programs collect follow-up data. Critics of RR argue that most follow-up studies use Clay's assessments to document long-term effects of the program. However, a number of state-wide follow-up studies conducted in the United States have utilized standardized measures to document that former RR children maintain their literacy gains and make average or better progress up to three, four and five years after the intervention ended.

Researchers at New York University tested a total of 1,596 second grade children and 604 third grade children who were successfully discontinued from RR between 1990-1993. The nationally standardized Slosson Oral Reading Test (1990) was administered to the total number of RR children and random sample children who participated in a follow-up study in the state of New York. The findings revealed that RR children's mean achievement levels on the Slosson word recognition test reflected average performance for students who were at the end of second grade, and slightly higher than average performance for those at the end of third grade, based on national norms. These results are impressive given that, only one and two years before, they were the lowest achieving students in their first grade classrooms.

Furthermore, 93% of the second graders and 98% of the third graders scored at or above grade level on a measure of text reading. The results of the four-year study demonstrate that the majority of the children in New York who had a full series of lessons and were successfully discontinued from RR in first grade sustained their gains and performed as well as their grade-level peers one and two years after completing the program. These results indicate that RR students in New York, after becoming average or better readers in first grade, continued to make significant progress in reading after the specialized teaching is discontinued (Jaggar, Smith-Burke, Ashdown, Simic, 1996).

A follow-up study conducted in the state of Massachusetts (RR Annual Report, 1996) produced similar results. In the Spring of 1995, 122 children who had successfully discontinued from RR during 1993-1994 and 143 non-RR children were randomly selected for a grade two study. The two groups of second grade children were compared on four measures: text reading, a story retelling, a dictation task, and the Slosson Oral Reading Test (1990). When compared with randomly selected non-RR students, the discontinued second grade RR students performed within the average band of achievement on text reading, story retelling, and the word recognition subtest on the Slosson. In May 1996, in addition to the same four measures, the Gates MacGinitie Reading Test (1989) was administered to the same groups of children who, at this time, were completing third grade. The achievement of discontinued RR students was compared to that of a random sample of third graders. The mean text reading level of 30.7 (roughly

equivalent to a grade 4 basal reader level) was achieved for former RR students and 31.0 for the random sample students providing evidence that both groups of students were reading well above grade level.

Furthermore, the former RR students, who were identified as the lowest achievers in grade one, were more successful at retelling stories than their random sample peers. Ninety-five percent of RR students and 92% of random sample students retold an end-of-grade-three story at an adequate to exceptional level. Scores from the Slosson Oral Reading Test (1990) and the Gates MacGinitie Comprehension Test (1989) demonstrated that the former RR students performed within the random sample's average band of achievement.

In a cross-sectional evaluation, researchers at Texas Woman's University (RR Texas State Report 1988-1996) studied second, third, and fourth graders who had successfully discontinued from Reading Recovery during their first grade year. Four assessments were used to measure the literacy performance of former RR students: a test of text reading, a written retelling, the comprehension test of the Gates MacGinitie Reading Tests (1987), and the reading subtest of the Texas Assessment of Academic Skills (TAAS). The results indicated that when compared to a random sample of their peers, former RR students placed well above grade level in text reading and written retelling and maintained their gains through fourth grade. Additionally, by fourth grade, former RR students compared well with their peers on the TAAS: 69% of the RR students and 76% of the random sample group of students had passing scores on TAAS. On the

Gates MacGinitie Comprehension Test (1989), 67% of the RR students and 71% of the random sample of fourth grade students had comprehension scores within the average range. These results indicate that former RR students are more similar to a random sample of peers on standardized measures when in grade four.

The follow-up studies from New York, Massachusetts, and Texas report on former RR students who were discontinued from RR; that is, these children reached the average reading levels of their peers and thus successfully completed the program in first grade. Due to limited resources, researchers in these three states could not follow every RR student who was served in the program in first grade. An Ohio fourth grade follow-up study, however, reports data on three groups of RR students: (a) those who were successfully discontinued from the RR program in grade one, (b) those who were referred for additional support, and (c) those who were served in RR but received fewer than 60 lessons (State of Ohio 1996-1997 Report).

The researchers in the Ohio study examined two cohorts of these three types of students: the first cohort received RR services in 1991-1992; the second cohort received RR services in 1992-1993. The children's overall proficiency scores on the Ohio Test of Fourth Grade Proficiency were examined for each cohort. The results revealed that when compared to all fourth grade children in the State of Ohio, 71% of the total group of RR students (including those who had fewer than 60 lessons, whether they were discontinued or not) in the first cohort scored above proficiency in reading and 72% scored above proficiency in writing. Similar results were reported for the sec-

ond cohort of children. Seventy-five percent of all RR students who were in first grade in 1992-1994 scored above proficiency on reading and 67% performed above proficiency on the writing measures. These findings suggest that the total number of RR students served made substantial gains in reading and writing by fourth grade.

When one considers that in every follow-up study, the random sample comparisons are drawn from a general population of regular education students who were not selected for RR and compares these students with former RR students who were once the lowest achievers in a first grade classroom, it is clear that the program does what it was designed to do—brings the hardest-to-teach children to a level of literacy achievement where they are full participants in classroom literacy programs. Furthermore, scores on two nationally standardized tests, the Slosson Oral Reading Test (1990) and the Gates MacGinitie Reading Test (1989), and on two statewide assessments, the Texas Assessment of Academic Skills and the Ohio Test of Fourth Grade Proficiency, collected at the end of grades two, three, and four, suggest that former RR students, in particular those that have been successfully discontinued (released) from the program, maintain their gains and continue to make reading and writing progress.

Conclusion

We are continually inundated with media reports that the number of school-aged children who cannot read and write continues to rise. Local, state, regional, and federal legislators, parents, business owners, and other stakeholders are demanding changes in school curricula

and practices to remedy the situation. In a comprehensive review of literacy programs in America's elementary schools, Allington and Walmsley (1995) argue that early reading achievement predicts future success or failure in life and strongly recommend early intervention programs for students who are failing to learn how to read.

Within this growing consensus for early intervention, there is continuing disagreement about several issues, including: (a) using professionals or paraprofessionals, (b) the amount of training necessary, (c) whether and how to monitor the fidelity of the local implementation of the program, (d) student-teacher ratio (e.g., one-to-one or small group), and (e) whether and how to collect data to document individuals' progress. Reading Recovery is decidedly consistent on all of these. Reading Recovery uses professionals who are required to be trained for a year, the pupil-teacher ratio is one-to-one, the local implementation is carefully monitored on a constant basis, and data are collected on individuals' reading and writing progress on a daily basis and reported annually to document the effectiveness of the program.

Any primary teacher can attest to the enormous range of differences in what children know and can do when they begin schooling. These individual differences suggest that the quantity, quality, and intensity of instruction needed to meet a child's idiosyncratic needs must differ. Allington and Walmsley (1995) encourage educators "to think of individual differences less as indicators of how much or how little children might learn, and instead think of them as indicating how much intensive instruction will be needed to accelerate their literacy devel-

opment and move them alongside their peers. As long as we believe that not all children can learn to read on schedule, we will fail to embrace instructional programs that demonstrate how wrong that tradition is” (p. 6). Perhaps that is what is happening in American elementary schools today—Reading Recovery supporters are challenging 100 years of conventional school organization and instructional practice.

In summary, RR operates within educational systems through three key program elements: (a) an intensive, daily, one-to-one, thirty minute program for the lowest achieving children in grade one; (b) an initial graduate level year-long training and continuous professional development program through which teachers refine their knowledge and skills in using proven techniques; and (c) a standard research program whereby individual data are collected on all students, even those who are served for one day, to monitor results continuously, to provide support for participating teachers, and to develop guidelines for implementing the program with integrity.

Reading Recovery has an approach to program evaluation that is coherent and which has employed both systematic and simultaneous replication extensively. I would encourage those who question RR to publish their program descriptions and their data along with replication information so that stakeholders will have substantive information for decision-making.

Essentially, after all of the objections to RR have been identified, after all of the arguments against the program have been weighed, we must face the hard and simple fact that no other program currently operating in the United States can pro-

duce thirteen years of data on every single child who was served in the program to document its success. At no time in recent history has there been more pressure to produce results. At no time in our history has there been a program that can produce more than a dozen years of replication data to document successful results.

References

- Anderson, R. C. (1988). *The Ohio Reading Recovery project: Preliminary report of the national evaluation panel*. Columbus, OH: Department of Education.
- Allington, R. L. (1994). What’s special about special programs for children who find learning to read difficult? *Journal of Reading Behavior, 26*, 95-115.
- Allington, R. L., & Walmsley, S. A. (1995). *No quick fix: Rethinking literacy programs in America’s elementary schools*. New York: Teachers College Press and International Reading Association.
- Armour-Thomas, E., Clay, C., Domanico, R., Bruno, K., & Allen, B. (1989). *An outlier study of elementary and middle schools in New York City: Final report*. New York: City Board of Education.
- Assad, S., & Condon, M. A. (1996). Demonstrating the cost effectiveness of Reading Recovery: Because it makes a difference. *Network News, 2*, 10-14.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand-McNally.
- Clay, M. M. (1993). *Reading Recovery: A guidebook for teachers in training*. Portsmouth, NH: Heinemann.
- Cunningham, P. M., & Allington, R. L. (1996). *Schools that work: Where all children read and write*. New York: Harper Collins.

- Darling-Hammond, L. (1996). The right to learn and the advancement of teaching: Research, policy, and practice for democratic education. *Educational Researcher, 25*, 5-18.
- Darling-Hammond, L., & Falk, B. (1997). Using standards and assessments to support student learning. *Phi Delta Kappan, 79*, 190-201.
- Dyer, P., & Binkney, R. (1995). Estimating cost effectiveness and educational outcomes: Retention, remediation, special education, and early intervention. In R. L. Allington, & S. A. Walmsley (Eds.), *No quick fix: Rethinking literacy programs in America’s elementary schools*. NY: Teachers College Press and the International Reading Association.
- Ferguson, R. E. (1991). Paying for public education: New evidence how and why money matters. *Harvard Journal of Legislation, 28*, 465-498.
- Frymier, J., Barber, L., Gansneder, B., & Robertson, N. (1989). Simultaneous replication: A technique for larger-scale research. *Phi Delta Kappan, 71*, 228-231.
- Gates & MacGinitie, (1989). *Gates MacGinitie Reading Tests (3rd ed.)*. New York: Columbia University and Teachers College Press.
- Gay, L. R. (1987). *Educational research: Competencies for analysis and application*. Columbus, OH: Merrill Publishing.
- Hersen, M., & Barlow, D. H. (1976). *Single case experimental designs: Strategies for studying behavior change*. New York: Pergamon.
- Hiebert, E. H. (1994). Reading Recovery in the United States: What difference does it make to an age cohort? *Educational Researcher, 23*, 15-25.
- Jaggar, A. M., Smith-Burke, M. T., Ashdown, J., & Simic, O. (1996). *A four-year follow-up study of Reading Recovery children in New York state: Preliminary report*. New York: New York University. *

- Kerslake, J. (1997). *A summary of the 1996 Data on Reading Recovery*. Research and Statistics Division Bulletin, No. 9. Wellington, NZ: Ministry of Education.
- Kratochwill, T. R. (1978). *Single subject research: Strategies for evaluating change*. New York: Academic Press.
- Learning disabilities—A barrier to literacy instruction*. (1995). Washington, DC: International Reading Association.
- Levin, H. (1989). Financing the education of at-risk students. *Educational Evaluation and Policy Analysis, 11*, 47-60.
- Lyons, C. A. (1994). Reading Recovery and learning disability: Issues, challenges and implications. *Literacy, Teaching and Learning: An International Journal of Early Literacy, 1*, 109-120.
- Lyons, C. A. (1997). Final report of the Reading Recovery program 1995-1996. Columbus, OH: The Ohio State University.
- Lyons, C. A., & Beaver, J. (1995). Reducing retention and learning disability placement through Reading Recovery: An educationally sound, cost-effective choice. In R. L. Allington & S. A. Walmsley (Eds.), *No quick fix: Rethinking literacy programs in America’s elementary schools* (pp. 116-136). NY: Teachers College Press and the International Reading Association.
- Massachusetts Superintendents Association Task Force. (1997). *The impact of special education on educational reform: Final report of the Task Force on Special Education*. Boston: The Massachusetts Association of School Superintendents.
- Moriarty, D. (1996). *Report to the Massachusetts superintendents task force on special education*. Medford, MA: Medford, MA Public Schools.
- Moriarty, D. (1997). A message to Congress: Redefining special education. *Network News, 1*, 16 -17.
- Neuman, S. B., & McCormick, S. (Eds.), *Single-subject experimental research: Applications for literacy*. Newark, DE: International Reading Association.

Reading Recovery in The United States

- Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. *Reading Research Quarterly*, 29, 8-39.
- Pinnell, G. S., Lyons, C. A., & Jones, N. (1996). Response to Hiebert: What difference does Reading Recovery make? *Educational Researcher*, 25, 23.
- Reading Recovery in Massachusetts (1996). *Technical State Report 1995-1996*. Cambridge, MA: Lesley College. *
- Reading Recovery in Ohio (1996). *Technical State Report 1995-1996*. Columbus, OH: The Ohio State University. *
- Reading Recovery in Texas (1996). *Technical State Report 1995-1996*. Denton, TX: Texas Woman's University.*
- Scott-Foresman and Company (1979). *Special Practice Books*.
- Shanahan, T., & Barr, R. (1995). Reading Recovery: An independent evaluation of the effects of an early intervention for at-risk learners. *Reading Research Quarterly*, 30, 958-997.
- Slosson, R.T. (1990). *Slosson Oral Reading Test* (3rd. ed.). East Aurora, NY: Slosson.
- Wasik, B. A., & Slavin, R. E. (1993). Preventing early reading failure with one-to-one tutoring: A review of five programs. *Reading Research Quarterly*, 28, 178-200.
- Wilson, K., & Daviss, B. (1994). *Redesigning education*, New York: Henry Holt.
- Woodcock Reading Mastery Test-Revised. (1989). New York: American Guidance Service.
- * follow-up studies

cal courses for teacher leaders and university trainers. Dr. Lyons has conducted research and published numerous articles in the field of cognitive processing, learning and reading disability, and teacher learning, curriculum development and practice. She is the coauthor of two books: *Partners in Learning: Teachers and Children in Reading Recovery* and *Bridges to Literacy*. Dr. Lyons is Past President of the Reading Recovery Council of North America.

Biography

Carol A. Lyons is a Professor of Education at The Ohio State University Columbus, OH. Dr. Lyons teaches graduate courses in developmental and corrective reading, literacy curriculum development and evaluation, teaching and learning, and the Reading Recovery theoreti-