DOCUMENT RESUME

ED 383 733 TM 023 173

TITLE	WCPSS Reading Recovery 1990-94. Evaluation Report. E&R Report No. 95.09A.
INSTITUTION	Wake County Public Schools System, Raleigh, NC. Dept. of Evaluation and Research.
PUB DATE	18 Apr 95
NOTE	48p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Early Intervention; *Educationally Disadvantaged; *Elementary School Students; Grade 1; Low Achievement; Outcomes of Education; Primary Education; Program Evaluation; Reading Difficulties; Reading Programs; *Tutoring
IDENTIFIERS	*Reading Recovery Projects; *Wake County Public School System NC

ABSTRACT

Reading Recovery (RR) is an early intervention program that provides intensive one-on-one services for first-grade students who experience difficulty in learning to read. RR teachers provide daily individualized lessons for 30 minutes until children are able to perform within the average range for their peers. RR was initiated in the Wake County Public School System (North Carolina) in the fall of 1990-91 and served 292 students in 35 schools in 1993-94. Evaluation has found that RR has not been able to meet its major short-term and long-term goals in Wake County. About half of all RR students have successfully reached first-grade reading levels, but success rates have declined since the first year of the program. Only one third of those who successfully reached grade level in 1990-91 and 1991-92 scored at grade level as third graders, about the same as comparison groups who received no RR services. No differences were found in needs for retention, special education, or Chapter 1 services. As currently implemented, RR is not enough to keep many Wake County low achievers reading at grade level. Eleven figures illustrate the findings. Two attachments and an appendix summarize results. (Contains 13 references.) (SLD)

Evaluation Report: WCPSS Reading Recovery 1990-94

April 1995

Wake County Public School System
Department of Evaluation and Research
E&R Report No. 95.09A

READING RECOVERY TERMINOLOGY

Categories used by Reading Recovery (RR) to describe specific groups of students served by the program can be difficult to understand. So that policy makers do not have to learn a whole new set of terms in order to understand program effects, we have simplified terms for RR subgroups as shown below. For the convenience of RR staff we have also provided RR "synonyms." Reading Recovery terms were taken from Pinnell, Lyons, and DeFord (1994).

Term Used	Synonym	Definition
Reading Recovery Students (all students served by RR)	Reading Recovery students	All students served who have received at least one RR lesson.
Full Service	Completers Program students	Students who received 60 or more lessons or who have been successfully discontinued.
▪ Successful (subgroup of Full Service)	▪ Discontinued ▪ Successfully Discontinued (subgroup of Completers)	Students who are able to successfully read at or above the average level in their class and have been released from the program.
▪ Not Successful (subgroup of Full Service)	▪ Not Discontinued (subgroup of Completers)	Students who have received 60 or more lessons but were not able to reach the average band for the first-grade reading levels.
Partial Service	Incomplete program	Students who received less than 60 lessons and were not successful in reaching the average first-grade reading band.

EVALUATION REPORT: WCPSS READING RECOVERY 1990-1994

TABLE OF CONTENTS

EVALUATION REPORT: WCPSS READING RECOVERY 1990-1994

REPORT SUMMARY

Reading Recovery (RR) is an early intervention program which provides intensive one-on-one services for first-grade students who experience difficulty learning to read. Reading Recovery teachers provide daily individualized lessons for 30 minutes until the children are able to perform within the average range for their first-grade peers; a full program is generally considered 60 lessons, although sometimes the number of lessons required will vary depending on students' progress. The program's major short-term goal is to help the lowest-achieving first graders become independent readers and writers. The program's major long-term goal is to lessen the need for subsequent special help through such means as retention, special education services, Chapter 1, or other programs.

Overall Results Summary

Reading Recovery has not been able to completely meet its major short- and long-term goals in Wake County Public School System (WCPSS). Results indicate that:

- About half of all RR students have successfully reached first-grade reading levels (two thirds of those receiving a full program), but success rates have declined across years (especially in 1993-94).
- Only one third of the 1990-91 and 1991-92 RR students who successfully reached first-grade reading level scored at grade level as third graders on the End-of-Grade (EOG) Reading test, about the same percent as the comparison groups who received no RR services.
- No differences were found in needs for retention, special education, or Chapter 1 two years later for 1990-91 RR and control students. Data from the 1990-91 and 1991-92 RR and comparison groups suggests some savings the first year after service.
- National and local cost data suggests that RR is very expensive relative to benefits seen.

Thus, it appears RR alone, as currently implemented, is not enough to keep many WCPSS low achievers reading at grade level. Of course, these analyses do not address the possibility that RR might have benefits that are not apparent until high school and beyond, as found for some pre-kindergarten programs.

Results have been organized in this report around four basic policy questions. A summary for each question follows, with more detail provided in the full report.

Q1. To what extent does Reading Recovery bring first graders who are having difficulty with reading up to grade level?

About half (372 or 48%) of the 772 WCPSS students served by RR in the last three years have successfully left the program reading at the first-grade level. Of those who received a *full* program, 67% were successful by this criteria. Over one fourth of the students served by RR did not receive a full program.

Reading Recovery's short-term success in WCPSS was consistent with national and state results in 1990-91, 1991-92, and 1992-93 (in terms of the percentage of students who finish the program reading at grade level), but dipped considerably below nationally reported rates in 1993-94. Success rates have declined over the four years RR has been in WCPSS.

Reading Recovery students who received full services in 1990-91 were more likely to reach average first-grade reading levels by the end of first grade than a control group of comparable students.

Q2. Does Reading Recovery impact students' need for further special assistance after first grade?

The evidence currently available indicates RR in WCPSS may have some positive impact on students' need for retention or Chapter 1 reading services in grade 2, but no impact by grade 3. Further one- and two-year follow-up studies are needed before firm conclusions can be drawn on cost savings. The evidence thus far does suggest RR alone is not enough for many students. At the student level, we found that:

- Reading Recovery students served in WCPSS in 1990-91, compared to a control group, were just as likely to be retained, placed in special education, or served by Chapter 1 in reading two years later. Results the year after service showed some inconsistent benefits for RR over comparison groups (with 1990-91 cohort less likely to need Chapter 1 and the 1991-92 cohort less likely to be retained).
- About *half* of the RR students served in 1990-91, 1991-92, and 1992-93 have needed further assistance through special education, Chapter 1, or retention. Among those who received a full program of services, those who were *successful* during the program were less likely to need further help than those who were not successful (28-34% of each successful cohort versus 55-79% of each not successful cohort).
- The percentage of RR students who needed further assistance did *not* decline for later cohorts as RR and school staff became more experienced with the program.

At the school level, first-grade retention rates have declined in WCPSS in both RR and non-RR schools. Reading Recovery may have been a contributing factor in reducing retention rates in RR schools.

Q3. Do students who are successful in Reading Recovery stay at grade level in reading after first grade?

Only about one third of the 1990-91 and 1991-92 RR students showed grade-level proficiency in third grade on the EOG Reading test. Successful RR students did not perform significantly better than those who were less successful, and RR students overall did not perform better than comparison students (who did not participate in RR).

Q4. Is Reading Recovery a cost-effective way to help our students who have the most difficulty learning to read succeed in school?

The average RR teacher serves seven students during a year, and, on average, three or four of those students read at a first-grade level by the end of the year. Annually, the cost per student for *all* students served in RR in WCPSS during 1993-94 was approximately $2,947.50 *beyond* the regular instructional program. The cost per *successful* student was approximately $6,000 beyond the regular instructional program. Current evaluation data suggests that by the end of third grade only about two of the students served by a RR teacher read at a third-grade level. Thus, the WCPSS has invested approximately $9,211 for each student who is a long-term success.

Since the 1990-91 and 1991-92 comparison groups of students who did not receive RR achieved a comparable success rate on standardized tests in third grade, and since RR expenditures in WCPSS do not seem to have been offset by significant savings from a reduction of need for special education, retention, or Chapter 1 assistance, the program does *not* appear to be cost effective at this time.

RECOMMENDATIONS

Early intervention programs like RR continue to hold promise, but the WCPSS program has not fully met its potential to date. Large-scale expansion of the program does not seem warranted at this time. In fact, based on national research, rapid expansion could diminish the chances for sound implementation and success. We recommend that the RR program be strengthened to optimize the chances for short- and long-term success. Whatever changes are made should be implemented in a way to allow a systematic study of their effects. Some of the critical issues raised by these research results are the depth of coverage in each school, the number of partial programs, possible follow-up with RR "graduates," and variations or alternatives to RR. A brief discussion of each issue follows.

The breadth and depth of RR coverage should be explored once again. When the program began, the decision was made to cover as many schools as possible, but not with great depth. In most cases, one teacher was provided per school. That decision meant that RR has not come close to the 15-20% coverage recommended by the originators of the program, and that the teachers have worked with more lower-achieving students than is true in districts with different policies. *Providing more teachers at fewer schools would allow a better test of the true potential for RR's success in WCPSS.*

The number of students receiving partial programs must be kept to a minimum, since they consume program resources but have a lower chance for success. The number of students with partial service could be reduced in a variety of ways. Some options to consider include finding ways to complete programs, perhaps through two lessons per day, or summer lessons, or lessons in the fall of second grade, and reducing the number of lessons teachers miss because of other duties.

Long-term results suggest some continuing support to former RR students (through classroom teachers, tutors, or Chapter 1) might be necessary to sustain reading at grade level. It is essential that the classroom teachers continue to build on the success students have in RR; RR teachers may be able to provide expanded training to primary-level teachers on how to accomplish this. In addition, periodic checks of RR graduate's skills and "booster" lessons as necessary are worth consideration.

While RR has been successful with some children, other alternatives or variations may be more cost effective or helpful to students long-term. New Title 1 guidelines for next year allow schools more flexibility in choosing ways to serve students. For example, Early Reading groups, a variation of RR using small groups, have been used thus far in WCPSS as a companion program (for those on a waiting list) and not an alternative to RR. Some schools may opt to use this less expensive group variation *instead* of RR, or to try RR in an after-school model (to avoid loss of instructional time for students). Alternatives will be included in a new publication coming from E&R called *What Works with Low Achievers? A Resource Guide*. If schools opt to implement a variation or alternative to RR, effects could be systematically compared to RR schools (especially those with at least 15% coverage) and no intervention at grade 1.

PROGRAM OVERVIEW

GENERAL

Reading Recovery (RR) is an early intervention program which provides intensive one-on-one services for first-grade children who experience difficulty learning to read. Reading Recovery teachers participate in a year-long training program which incorporates peer support, group discussion, coaching by experienced teachers, and reflection time (Pinnell, Lyons, DeFord, Bryk, and Seltzer, 1994). Reading Recovery teachers provide daily individualized lessons for 30 minutes until the children are able to perform within the average range for their first-grade peers; a full program is generally considered 60 lessons, although sometimes the number of lessons required will vary depending on students' progress (12-20 weeks generally). Within each 30-minute lesson, the students are engaged in a variety of instructional activities such as: reading and rereading books, writing and reading their own stories, and analyzing letters and sounds in words. Children are taught to use many strategies to understand printed text and pictures (just as successful readers use). In addition to direct work with students, RR staff provide consultation and training to school staff about reading instruction.

The program's major short-term goal is to help the lowest-achieving first graders become independent readers and writers. The program's major long-term goal is to lessen the need for subsequent specia. help through such means as retention, special education services, Chapter 1, or other programs.

NEW ZEALAND VERSUS UNITED STATES AND WCPSS READING RECOVERY PROGRAMS

Marie Clay founded the RR program in New Zealand in 1979. Clay observed program implementation and made refinements over a number of years to enhance success. Regular first-grade teachers taught their lowest readers one-on-one in a RR model for half of the day and their full class the other half of the day. Approximately the lowest 20% of the students in terms of reading skills were served, and the program was very successful, bringing 90% of these students to average first-grade reading levels and keeping them at that level without further assistance (only 0.8% of successful RR students needed further help).[1] The program was first implemented in the U.S. in the mid-1980's in Ohio Public Schools in conjunction with Ohio State University. The National Diffusion Network (NDN) selected RR as an exemplary program in 1987 based on those positive results. Figure P1 (page 4) provides more information on NDN and their selection process.

---

[1] Figure presented by Clay at the Reading Recovery Conference on Implementation in Durham, December 1994.

One difference between New Zealand and the United States (U.S.), at least at present, is that *all RR students receive a full program in New Zealand*, which is not true here. Based on 1990-91 and 1991-92 data, about half of the Columbus, Ohio RR students received partial programs (Pollock, 1991 and 1993). (In WCPSS, at least 25% of the students received a partial program, rather than a full program of 60 or more lessons.) Partial programs are less common in New Zealand for two basic reasons:

- *The school year is longer (220 days).* This allows more programs of 60 lessons to be completed (at least three groups of four students per year rather than two groups of four in the U.S.). If a program is not completed, the rest of the program is generally completed at the beginning of grade 2 after the much shorter summer break.
- *Each school in New Zealand has "full coverage" (i.e., adequate RR teachers to reach the lowest 20% of their students).* Therefore, if students transfer, they are picked up by the program at the new school. WCPSS currently has coverage for 5-7% of the students in the schools with the program in 1993-94 (or 4.4% of the first graders in WCPSS overall).

Another key difference is that the regular classroom teacher is also the RR teacher in New Zealand. This offers an opportunity to reinforce learning in the regular classroom, and this is not present in the models more commonly used in the U.S. According to Clay, this was originally done because the RR teachers believed they would miss the regular classroom environment (not because of the intensity of the work). The models used in the U.S. vary, with some school districts hiring teachers who only work half-time and some hiring teachers who teach RR half-time and have other duties half-time. In WCPSS, RR teachers spend the other half of their time either teaching Early Reading to small groups (students on the waiting list for RR) or consulting with other teachers in the school.

Finally, New Zealand does not recognize some of the special education categories we use in the U.S. (such as Learning Disabled) and generally tends to refer students a little later for special education (age 7 on). It is therefore difficult to compare special education rates in the U.S. with those in New Zealand, and it is less likely that a student in New Zealand would be referred to special education before RR had a chance to be completed with a child.

HISTORY OF READING RECOVERY IN WCPSS

Reading Recovery was first implemented in the WCPSS in the fall of 1990-91 in 11 schools. The number of schools involved has tripled since that time. The number of schools and teachers involved each declined by one between 1992-93 and 1993-94, while the number of students served increased. Enrollment for 1990-91 through 1993-94 is shown below.

Year	Number of Schools	Number of Teachers	Number of Students	Average Students per Teacher
1990-91	11	12	84	7.0
1991-92	20	24	146	6.1
1992-93	36	42	250	6.0
1993-94	35	41	292	7.1

The major cost of the RR program is for staffing. One trained half-time RR teacher serves four students at a time, meeting individually with each student for daily half-hour sessions for periods ranging from approximately 12 to 16 weeks. Under optimum conditions, a half-time teacher could serve 8-12 students per year, but because of a variety of factors, including time for testing and selecting of students, as well as student and/or teacher absences, RR teachers in WCPSS have averaged only 6-7 students each during the past four years. Dividing the cost for one half of an average teacher's salary plus benefits in 1993-94 ($20,632.50) by seven yields a per-child-served cost of $2,947.50 in addition to the cost of the regular instructional program. (See page 25 for further discussion of costs and outcomes.) Most of the schools are funded through Chapter 1, with local money funding the remainder of schools. In 1993-94, six of the 35 schools with RR service had local funding.

Figure P1. About NDN Programs

What is the National Diffusion Network (NDN)?

NDN is one mechanism through which the U.S. Department of Education identifies and publicizes exemplary programs. Annually, NDN publishes a compendium of recognized programs. For a program to be selected as exemplary, short applications (up to 15 pages) containing data on the program's effectiveness must be submitted. Each is reviewed by a subcommittee of six members of the 60-member Program Effectiveness Panel (PEP).

What criteria do PEP panelists use in evaluating applications to NDN?

The six panelists rate the applications on three dimensions: results, evaluation design, and replication. Ratings focus primarily on program outcomes. The largest weight is given to the results dimension (maximum of 50 points), with heavy weight also given to the description of the evaluation design (up to 40 points). Longitudinal or long-term results are not required, but quantitative studies have historically been rated higher than qualitative studies. The replication dimension accounts for 10 of the maximum of 100 possible points. A median total score of 70 points or higher is required for selection.

Does the NDN process guarantee that programs selected as exemplary will work in another school or school system?

Not necessarily. Although NDN is a good source of initial information, the NDN review process has recently come under fire. An overview of the limitations of the NDN process can help practitioners better understand what NDN endorsement means and does not mean. First, the process pays little attention to each program's goals, unique or essential features, training needs, etc. Reviewers may know very little about the essential details of a program or its implementation when assigning their ratings. No site visits are made, nor are students, teachers, or parents contacted. The review process also has these additional limitations:

- NDN "does little to help potential customers understand the relative strengths and weaknesses of the programs intended to accomplish similar goals" (Klein, 1993);
- Flaws in research design or data collection may be undetected in the review process;
- Panel reviews of the applications "can do little to uncover misleading claims" (Walberg and Niemiec, 1993); and
- "Marginal or un-meritorious programs may win panel approval" if the application is prepared by someone who writes well and is familiar with the rating scale. (Walberg and Niemiec, 1993.)

Finally, for program results to be replicated in a new site, it is often the case that the programs must be implemented in exactly the same way as they were in the successful sites. Differences in school structures and operations can make a difference. This exact duplication is not always feasible, and even when feasible, variations in program implementation can occur for a myriad of reasons.

EVALUATION MODELS

This report focuses on four issues of key interest to policy makers. Given the high number of priority projects on E&R's agenda for the year and our limited resources, we focused on short- and long-term student outcomes, which we see as the key to overall program effectiveness. The program staff already collects their own information on program implementation and use pre- and posttest results from the Clay Observation Survey in monitoring their short-term success; we have incorporated some of their key results into this report to provide a fuller view of overall program impact on students.

This evaluation includes student subgroup results of particular value to program staff, but was not designed to give staff broad-based feedback on the success of their professional development efforts with school staff (except as indirectly reflected in student outcomes). In addition, the report does not focus directly on staff or parent attitudes towards RR. Views of both these groups have been quite positive. However, in our opinion, the ultimate effectiveness of RR must be based on student outcomes.

Information on how the four RR policy questions are addressed in this report is presented below.

Q1. To what extent does Reading Recovery bring first graders who are having difficulty with reading up to grade level?

This question is addressed in the Short-Term Results section of this report. E&R examined the percentage of the students served in each RR group who reached the first-grade level by the end of the first grade. To help district policy makers with budget decisions, all students served (including those with partial programs) were included in the E&R analysis, since resources were spent for all of these students. This provides the best perspective on whether RR is cost-effective.

Data reported to the National Diffusion Network (NDN) by RR staff focused **only** on those RR students *who received a full program and who were successful in reaching first-grade reading levels*. In the NDN model, "first-grade level" was based on the performance of a random sample of first graders in WCPSS on the Clay Observation Survey. Focusing on those receiving full service by RR is very useful to program staff in determining whether their program has been successful when implemented fully with students. It does not, however, address the broader policy issue of the overall value or impact of RR services on all students served (including those who did not receive full services). In WCPSS, 25-34% of RR students have received only a partial program; rates from other studies have been even higher (Pollock 1991, 1993).

The impact of the two reporting methods can best be illustrated through an example. The 1991-92 report on Columbus, Ohio (Pollock, 1993) indicates that 227 students were served and 117 (52%) received a full program. Of this 117, seventy-eight were successful in reaching the average band for grade 1. If success rates are computed based on all students

who received a *full* program, 67% were successful; if all students served *fully* or *partially* are included, only half as many (34%) were successful. Systematically excluding those who receive partial programs from success analyses thus hides a substantial cost and serves to inflate overall success figures reported.

Some comparison data is reported for North Carolina and studies in other sites in the U.S. and New Zealand (primarily for those students who received full services).

Q2. Does Reading Recovery impact students' need for further special assistance after first grade?

In the Long-Term Needs For Further Assistance section, we examined the impact of RR on students' long-term need for subsequent special assistance through special education, Chapter 1, or retention the year after service. E&R staff also investigated whether first-grade retention rates were lower overall at the school after RR program implementation.

In addition, we explored whether the need for special assistance declined as RR staff became more experienced. To determine whether changes were due to RR, we examined the extent to which RR students needed special help compared to a very sound control group *of similar students in the same schools* but not in the program. E&R analyses included RR students served overall and by subgroup to allow examination of further service needs. NDN does not require this long-term look at program impact. Studies done at other sites in the U.S. were available for limited comparisons, but research designs varied. New Zealand rates for further assistance were of limited value for direct comparison because of the differences in special education and other available programs.

Q3. Do students who are successful in Reading Recovery stay at grade level in reading after first grade?

The section Long-Term Results on Reading Achievement addresses this question. E&R staff looked at test scores on various measures for students in grades 3 and 4. Generally, the students who were successful in first grade in RR were those of key interest for this specific question, although the subsequent reading achievement of those who were less successful in the program provided useful context and comparison data to help in interpreting results. Local results and some from other sites were available.

Q4. Is Reading Recovery a cost-effective way to help our students who have the most difficulty learning to read succeed in school?

Reading Recovery is expensive initially because teachers work with students individually. However, RR program advocates argue that these costs are lessened because students are more likely to be successful in subsequent years without additional help (through Chapter 1, special education, or retention). In the Cost-Effectiveness section, information on monetary costs is combined with the short- and long-term results to address the overall question of cost-effectiveness.

SHORT-TERM RESULTS

Q1. To what extent does Reading Recovery bring first graders who are having difficulty with reading up to grade level?

About half (372 or 48%) of the 772 WCPSS students served by RR in the last three years have successfully left the program reading at the first-grade level. Of those who received a *full* program, 67% were successful by this criteria. Over one fourth of the students served by RR did not receive a full program.

Reading Recovery's short-term success in WCPSS was consistent with national and state results in 1990-91, 1991-92, and 1992-93 (in terms of the percentage of students who finish the program reading at grade level), but dipped considerably below nationally reported rates in 1993-94. Success rates have declined over the four years RR has been in WCPSS.

Reading Recovery students who received full services in 1990-91 were more likely to reach average first-grade reading levels by the end of first grade than a control group of comparable students.

PROCEDURES AND LIMITATIONS

The research design RR utilizes for the National Diffusion Network (NDN) involves determining the extent to which students served in RR are able to meet a band of average reading performance for first graders in the school system. This school system standard is based on a random sample of WCPSS first-grade students tested with the Clay Observation Survey in the areas of Text Reading Level, Dictation, and Writing Vocabulary. (See Figure S1 for more information on the Clay Observation Survey.) Reading Recovery students are considered successful and "discontinued" from services if they demonstrate reading skills within a band of 0.5 standard deviations above or below the means for this group.

At the end of each program year, the RR staff summarizes group trends for the Clay Observation Survey in terms of the extent to which RR students were able to "close the gap" between their performance and the average band for the random group of WCPSS first graders. The main difference between results as reported to NDN and those reported in E&R reports is that reports to NDN do not include any students in the analyses who do not receive a full program, which eliminates from their analyses about 25% of the students served.

WCPSS extended the NDN evaluation model by including comparison groups of similar low achievers who were not served by RR. In 1990-91, a control group was established within the same RR schools by randomly assigning half of the eligible students to RR and half to a control group. The fact that control students were in the same schools and of equal need to those served made this a strong control group. However, the small size of the program in

1990-91 and the fact that 1990-91 was the first year of the program (which meant teachers were less experienced) means caution is advisable in drawing conclusions. The teacher experience issue seems to be a minor concern since discontinuation rates were actually highest that first year of RR.

In 1991-92, a comparison group was developed which represented the lowest readers in non-RR schools (one school which was eligible and declined participation and the other schools were just above the cut-off for eligibility). The Clay Observation Survey was administered to all non-program comparison students to allow comparisons of their performance to that of RR students. After E&R's initial report was published on needs for further assistance (Donley and Baenen, 1993), the RR staff raised a concern that the groups might not be comparable because they were not able to serve as high a percentage of the students in the RR schools as were represented in the control group. A check of pretest scores revealed that the comparison group did, in fact, score higher than the RR students on the pretest. For this reason, we did not intend to report results for the 1991-92 RR cohort and their comparison group. However, in the review process, the RR staff requested that these results be included, because the RR students scored higher on the spring Clay Observation Survey than the comparison group. While we still have design concerns about the initial comparability of the groups, we have included results as requested.

Figure S1. What is the Clay Observation Survey?

Marie Clay developed the Clay Observation Survey to select the lowest readers for RR and assess their progress after service. It includes six measures. (Pinnell, Lyons, & DeFord, 1994.)

- **Letter identification:** Students identify upper- and lower-case letters.
- **Word Test:** Students read 20 common words.
- **Concepts about Print:** Students perform different tasks (i.e., directionality and word concepts) related to printed language concepts while reading a book.
- **Writing Vocabulary:** Students write all the words they know in 10 minutes (spelling counts).
- **Dictation Test:** Students write the words that are dictated to them in a sentence form.
- **Text Reading Level:** Students read until their accuracy rate falls below 90%. (Text difficulty is drawn from a basal reading system that is not part of the RR instruction.)

RESULTS

Reading Recovery Cohorts From 1990-91 Through 1993-94

The major short-term goal of the RR program is to help the lowest readers in a first-grade cohort reach the average reading level range of their peers by the end of first grade. Students selected for service in RR can be divided into three groups at the end of first grade:

- Students who received the full program and successfully reached an average reading level (usually requiring 60-70 sessions, but sometimes accomplished in fewer sessions);
- Students who received the full program (at least 60 tutoring sessions) and did not successfully reach an average reading level; and
- Students who received a partial program (less than 60 tutoring sessions) and did not reach an average reading level.

As shown in Figure S2, the *number* of RR students served and the number successful increased as the number of teachers involved increased. However, only half of those served overall (48%) successfully reached the average first-grade reading range.

Figure S2. Teachers and Student Enrollment in WCPSS Reading Recovery Program

Year	Teachers	Total Students Served	Full Program			Partial Program
			Successful	Unsuccessful	Total	
1990-91	12	84	47 (56%)	14 (17%)	61	23 (27%)
1991-92	24	146	77 (53%)	28 (19%)	105	41 (28%)
1992-93	42	250	122 (49%)	44 (18%)	166	84 (34%)
1993-94	41	292	126 (43%)	94 (32%)	220	72 (25%)
TOTAL	**119**	**772**	**372 (48%)**	**180 (23%)**	**552**	**220 (28%)**

Note: Teachers provided Reading Recovery half-time

As shown in Figure S2, at least one fourth of the students do not receive a full RR program by the end of the first-grade school year, and a significant investment of a RR teacher's time is devoted to these students. An analysis of the 84 students who received only a partial program in 1992-93 reveals that:

- Fifty-six (67%) were still in the program with less than 60 lessons at year's end,
- Seventeen (20%) were withdrawn to a special education program with most (11) classified as Learning Disabled,
- Six (7%) transferred from a RR school before receiving 60 lessons,
- Five (6%) withdrew for other reasons from RR with less than 60 lessons.

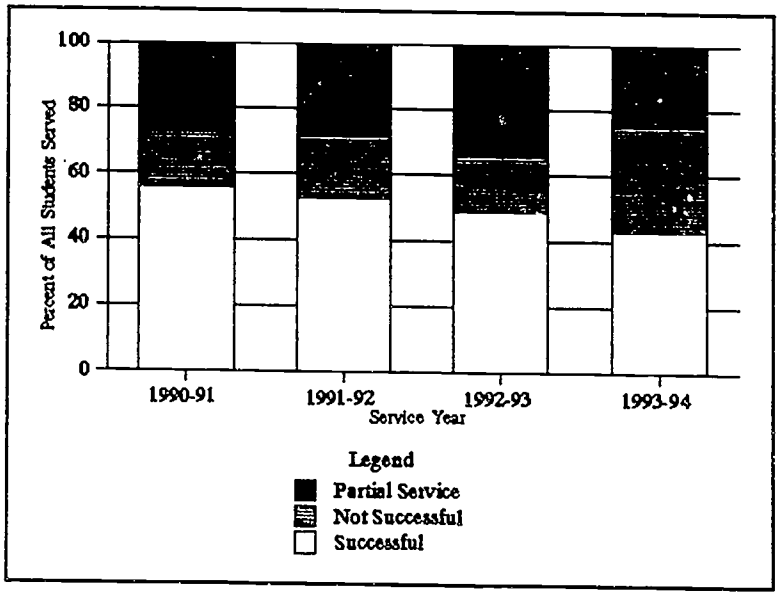
Thus, two thirds of the students receiving partial programs simply are phased into the program too late to have time to receive 60 lessons before the school year ends. This occurs because teachers work with only four students at a time and add new students as others become successful and leave the program. This suggests a need for better efforts to provide full services for more students.

Among those students who received a full program, the percentage able to successfully reach the first-grade average level has been:

Year	Full program students who reached first-grade average level
1990-91	77%
1991-92	73%
1992-93	73%
1993-94	57%
TOTAL	67%

When looking across years, approximately 67% of the 552 students who received a full program were successful in reaching the average first-grade reading level.

Figure S3. Completion Status for All Students in Reading Recovery from 1990-91 to 1993-94



As shown in Figures S2 and S3, *there has been a statistically significant downward trend across years in the percentage of all students served who successfully reach an average first-grade reading level*, dropping from 56% in 1990-91 to 43% in 1993-94. The decline is similar (and significant) for those receiving a full program--declining from 77% to 57%. *Much of this decline was seen in 1993-94.* In contrast, the National Data Evaluation Center reported that in 1993-94, 82% of the children in North America who received a full RR program were successful. (Pinnell et al., 1994.)

In North Carolina (Jones, 1994), the 1991-92 and 1992-93 discontinuation rates for all students served and those who received full services were consistent with rates found in WCPSS (see Figure S4). However, Jones notes that North Carolina discontinuation rates were lower than those found in RR nationally in those years (67-78% for those receiving full programs) and that the number of weeks to successfully discontinue students has been higher (18 weeks on the average) than in other states.

Figure S4: North Carolina and WCPSS Percentage of Successful Students

Students	Year	NC	WCPSS
All Served	1991-92	52%	53%
	1992-93	45%	49%
Full Service Only	1991-92	72%	73%
	1992-93	68%	73%

Jones suggests three factors that may play a part in these trends statewide: rapid expansion into varying types of sites, teachers new to the program, and less than full coverage (i.e., insufficient teachers to reach the lowest 15-20% of readers).

One reason for these lower success rates in WCPSS may be that the RR program is being provided only to the lowest 5-7% of readers in each school using the program. Studies at other U.S. sites of RR have shown higher success rates, but in those locations the program is reaching more students, often approaching 15-20% coverage (as recommended by RR literature). Increasing coverage would mean providing services to students having slightly less difficulty; this may increase the overall success rate for the program, although not necessarily improving the success rate for the lowest achieving students. Reading Recovery program staff believe absenteeism may have been a factor in the 1993-94 dip in success rates. In 1994-95 they are addressing that issue, as well as ways to teach the most challenging students more successfully.

Reading Recovery Students Versus Comparison Students

A comparison of short-term results on the Clay Observation Survey for the 1990-91 RR students who received full services and the control group found positive patterns.

- Reading Recovery students showed greater short-term gains than control students in Writing Vocabulary, Dictation, and Text Reading;
- A higher percentage of RR students scored in the first-grade average band than the control group (80% vs. 45% for Writing Vocabulary, 61% vs. 35% for Dictation, and 49% vs. 15% for Text Reading).

A comparison of short-term results on the Clay Observation Survey for the 1991-92 RR students who received full services and the comparison group also revealed positive patterns. Reading Recovery's full-service students had lower pretest scores in Writing Vocabulary, Dictation, and Text Reading, but higher spring posttest scores in all three areas.

LONG-TERM NEEDS FOR FURTHER ASSISTANCE

Q2. Does Reading Recovery impact students' need for further special assistance after first grade?

The evidence currently available indicates that RR in WCPSS may have some positive impact on students' need for retention of Chapter 1 reading services in grade 2, but no impact at grade 3. Further one- and two-year follow-up studies are needed before firm conclusions can be drawn on cost savings. The evidence thus far does suggest RR alone is not enough for many students. At the student level, we found that:

- Reading Recovery students served in WCPSS in 1990-91, compared to a control group, were just as likely to be retained, placed in special education, or served by Chapter 1 in reading two years later. Results the year after service showed some inconsistent benefits for RR over comparison groups (with the 1990-91 cohort less likely to need Chapter 1 and the 1991-92 cohort less likely to be retained).
- About *half* of the RR students served in 1990-91, 1991-92, and 1992-93 have needed further assistance through special education, Chapter 1, or retention. Among those who received a full program of services, those who were *successful* during the program were less likely to need further help than those who were not successful (28-34% of each successful cohort versus 55-79% of each not successful cohort).
- The percentage of RR students who needed further assistance did *not* decline for later cohorts as RR and school staff became more experienced with the program.

At the school level, first-grade retention rates have declined in WCPSS in both RR and non-RR schools. Reading Recovery may have been a contributing factor in reducing retention rates in RR schools.

PROCEDURES AND LIMITATIONS

To address this policy question, E&R conducted both school-level and student-level analyses for the 1990-91, 1991-92, and 1992-93 RR cohorts.

SCHOOL-LEVEL ANALYSES

One way to think about this issue is whether RR impacted the extent to which *schools* needed to provide future assistance to students. For students without a control or comparison group, it is not possible to estimate the extent to which the students actually served in RR would have needed additional help. Reading Recovery sites which have studied the long-term impact have sometimes checked overall retention, special education, and Chapter 1 service needs at the *school* level before and after RR. One limitation of this approach is that it makes the assumption that RR will have an impact on statistics covering all first graders in a

school even though RR provides service to a maximum of 20% of the students. Another limitation to this approach is that evaluators cannot control for other factors, in addition to RR, which impact retention rates. Finally, this design compares results for two successive groups of different first graders, before and after RR implementation, who may have varying levels of need.

WCPSS program staff perceived that RR had impacted school retention rates. We therefore checked retention rates at the first-grade level the year before RR and one and two years after implementation. The specific year of the program implementation varied across schools because RR was phased in; the program began in eleven schools in 1990-91 and expanded in subsequent years to a total of 35 schools in 1993-94. Retention rates for individual schools and for the district were taken from Long-Range Planning Department reports (a report which was not available for 1993-94). We were able to compare rates in 30 schools after the first year of the program and in a subset of 17 schools after the first and second year. Schools which had no students retained at first grade during each of the two years being compared were not included in the analysis. We checked these changes against district trends for the same time period to help determine to what extent RR could be credited with changes seen.

STUDENT-LEVEL ANALYSES

A second way to explore this issue is based on RR *students* across time and in comparison to a control group. This model provides more control for the impact of RR (versus other factors). We were not able to find any RR studies at other sites which employed this model. We checked the 1990-91 RR and control student records one and two years after they received services to determine whether they were receiving special education, retention, or Chapter 1 reading services. We also checked 1991-92 RR and comparison group students for further assistance needs one year after service. (The 1991-92 comparison group is not called a control group because they were not as closely matched as in 1990-91.) These analyses addressed the following evaluation questions:

- Were RR students less likely to need further assistance than the 1990-91 and 1991-92 comparison groups?
- What percentage of RR students needed further assistance overall?
- Were later cohorts of RR students less likely to need further assistance (after RR staff were more experienced and had trained regular classroom teachers more)?
- Were students who were successful in RR less likely to need further service than those who were less successful?

The extent to which RR students need further assistance is helpful in determining long-term costs and in checking the assumption that part of RR's initial cost can be recovered because students are less likely to need other special help in future years.

Students' need for any type of special education services was examined, as well the number specifically categorized as Learning Disabled (a group for which RR expects the greatest

impact). One limitation is that we do not have the subject area of service (reading versus mathematics or another area) available at the present time because it is not listed on mainframe computer files. (Reading Recovery staff are obtaining that information from the schools at the present time.)

An additional point to keep in mind in interpreting results by subgroup is that student records were not specific enough, nor sample sizes large enough, to separate students who received Early Reading service prior to RR from those who did not. Among those eligible for RR, those with relatively less need (higher pretest scores) were served through Early Reading (a group approach based on RR principles and strategies) until space became available in RR. These students may therefore be more common in the partial-service group. This clouds the expectations somewhat for the differences expected between full-service students and partial-service students (i.e., full-service students may not show higher posttest scores or lower needs for further assistance than partial-service students).

RESULTS

SCHOOL-LEVEL RESULTS

Several school districts have reported decreases in first-grade retention rates and special education placements after implementing RR. For example, RR schools in the Wareham School District in Massachusetts reported reduced first-grade retentions in their RR schools from 14 children to zero and special education placements from 31 to 16 after the school's first year of RR service in 1990-91 (Dyer, 1992).

Reading Recovery, along with other factors, may have contributed to lower school retention rates for first graders in WCPSS. Even before RR, however, first-grade retention rates were quite low in the schools involved (2.8%). Rates declined even further after RR was implemented, as shown in Figure L1 (see Attachment 1 for results by school). It is important to note that districtwide trends were continually decreasing during the years 1989-90 to 1992-93, so RR might have been one of several contributing factors in the decrease of retention rates.

- Overall, retention rates decreased in two thirds of the schools the first year RR was implemented compared to the year before the program, while one third of the schools had an increase in retention. The schools' average retention rate the year before RR was cut in half (from 2.8% to 1.4%) the first year RR was used.
- During the second year of RR, average retention rates remained lower than the year before RR was started (from 2.7% to 1.7%). Most schools (12 out of 17) had lower retention rates than the year before the program, while a few schools (5 out of 17) had higher rates.

Figure L1. School-Level Average Retention Rates for Grade 1

Group	Number of Schools	Before RR	1 Year After RR	2 Years After RR
One-Year Comparison	30	2.8%	1.4%	NA
Two-Year Comparison	17	2.7%	1.3%	1.7%

STUDENT-LEVEL RESULTS

Reading Recovery and Control Group 1990-91

In the year following service, RR students served in 1990-91 *were not* less likely than a control group of very similar students to need assistance in the form of *special education* or *retention*, but *were* less likely to need *Chapter 1* service. ***Two years after RR services, no differences were found in assistance needs between the two groups.*** (See Figure L2.)

- Reading Recovery students were about as likely as the control group students to be retained by the fall of 1992 when they should have been entering third grade (9.7% versus 9.3%, respectively). About 5% of each group were retained in grade 1 and 4% in grade 2.
- Reading Recovery students were about as likely as the control group to be receiving special education services one year later (13% versus 13%) and two years later (11% versus 8%). Patterns were the same for LD placements (3% versus 5% after one year and 6% versus 7% after two years).
- Reading Recovery students were significantly less likely than control group students to receive Chapter 1 reading services the *next* fall (32% versus 51%) but *not less likely* to need service by two years later (35% versus 37%).
- Reading Recovery students were significantly less likely than the control group students to receive any further assistance the next fall (43% versus 63%) but *not* by two years later (46% versus 48%).

Thus, when all RR students were compared to the control group, no differences were found two years later. *When just those who received a full program from RR were compared to the control group, the pattern was the same*, with no significant differences found two years later between proportions served by retention (4% versus 4%), special education (4% versus 8%), Chapter 1 (33% versus 37%), or any service (37% versus 48%).

Reading Recovery and Comparison Group 1991-92

Data were available to compare the 1991-92 RR and comparison groups one year following service. Based on all students served by RR, the following trends were evident one year later: (See also Figure L2.)

- Reading Recovery students were significantly less likely to be retained in grade 1 than the comparison group (8% versus 21%).
- Reading Recovery students were about as likely as the comparison group to be receiving special education services (23% versus 25%). The same pattern was true for LD placements (16% versus 10%).
- Reading Recovery students were about as likely as the comparison group to receive Chapter 1 services (29% versus 31%).
- Reading Recovery students were about as likely as the comparison group to receive any service (54% versus 64%).

When just those who received a full program from RR were compared to the comparison group, the pattern was similar.

Figure L2. Percent of Reading Recovery and Comparison Groups Needing Further Assistance One and Two Years Later

One Year Later

Cohort	Group	Retention	Special Education	Chapter 1	Any Assistance
1990-91	All Reading Recovery (n=72)	4 (6%)	9 (13%)	23 (32%)	31 (43%)
	Full Service (n=51)	1 (2%)	3 (6%)	17 (33%)	19 (37%)
	Control Group (n=75)	4 (5%)	10 (13%)	38 (51%)	47 (63%)
1991-92	All Reading Recovery (n=141)	11 (8%)	32 (23%)	41 (29%)	76 (54%)
	Full Service (n=101)	4 (4%)	22 (22%)	29 (29%)	49 (49%)
	Comparison Group (n=87)	18 (21%)	22 (25%)	27 (31%)	56 (64%)

Note: Shaded boxes showed significantly smaller proportions served than the control group.

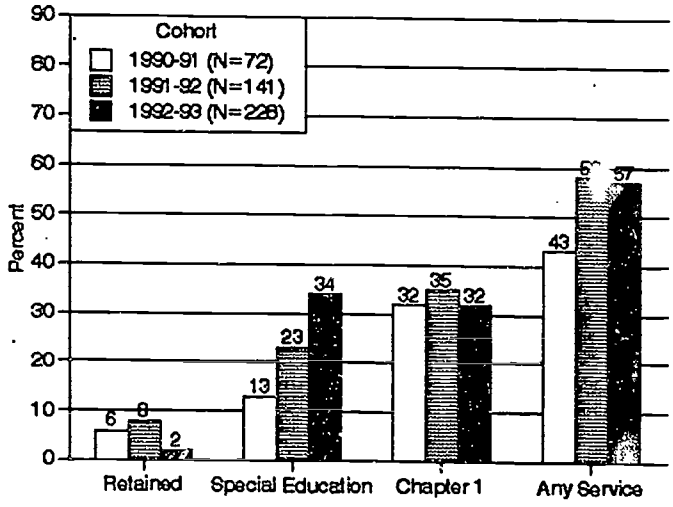
Two Years Later

Year	Group	Retention	Special Education	Chapter 1	Any Assistance
1990-91	All Reading Recovery (n=72)	3 (4%)	8 (11%)	25 (35%)	33 (46%)
	Full Service (n=51)	2 (4%)	2 (4%)	17 (33%)	19 (37%)
	Control Group (n=75)	3 (4%)	6 (8%)	28 (37%)	36 (48%)

Reading Recovery Students' Need for Further Assistance Across Years

<u>All Reading Recovery Students</u>. *Overall, 43%-58% of the RR students served in 1990-91, 1991-92, and 1992-93 received further service through retention, special education, and/or Chapter 1 the following year.* When Learning Disabled was the only special education service considered, this percentage dropped slightly, to 42%-53%. (See Figure L4.)

Figure L4. All Reading Recovery Students: One Year Follow-Up



Sub-Group Results. Patterns of assistance needs were similar for each of the subgroups studied (full service, partial service, successful, and not successful RR students). Additional analyses comparing full to partial service and successful to not successful students found two positive patterns worthy of note. (For more details, see Appendix A.)

- *Reading Recovery students who received full services were less likely than those who received partial services to need some form of subsequent assistance for two of the three groups studied (the 1991-92 and 1992-93 groups but not the 1990-91).* About one half of the students receiving full services through RR in 1991-92 and 1992-93 needed subsequent help versus three fourth's of those receiving partial service.

  When examined by types of assistance, the pattern was inconsistent, with only four of nine comparisons favoring students who received full services. The pattern remained similar when only special education placement into the Learning Disabled category was considered.

- *Within those who received a full RR program, successful students were significantly less likely to need some type of subsequent assistance than not successful ones in all three cohorts.* Among successful RR students, 28-34% needed subsequent assistance, and among the unsuccessful, 55-79% needed further assistance.

  When examined by type of assistance, successful students were significantly less likely to be retained in grade or placed in special education or Chapter 1 in only four of nine comparisons. Patterns were not consistent across type of service. Patterns for 1991-92 and 1992-93 *were* more positive than in 1990-91.

Comparison Across Reading Recovery Cohorts.

It was hoped that the need for further assistance after RR would decline as RR and other staff became more experienced with the program. However, when overall subsequent need for service was considered, *students served by RR in 1992-93 were not less likely to need subsequent assistance than the groups served in 1991-92 and 1990-91.*

When types of assistance were considered separately, RR students served in 1992-93 were *more* likely to be placed in *special education* (34% vs. 19%). Rates for referral to the Learning Disabled category were lower than for special education overall (4-16% across years compared to 13-34%). The 1992-93 RR group was not less likely to be placed in the Learning Disabled category as the 1991-92 group (16%) and both groups exceeded the 4% rate found for the 1990-91 RR students. The 1992-93 RR students were *less* likely to be *retained* (2% vs. 7%) than the groups served earlier. There was no significant difference in the percentage served by Chapter 1 (32% vs. 34%).

Those who received full RR services in 1991-92 and 1992-93 were less likely to need some type of additional assistance than those in 1990-91. In addition, within the full-service group, those successful in reaching the desired first-grade reading range were less likely to need additional follow-up services than those who were less successful. These patterns by subgroup are encouraging, although patterns within specific types of service are inconsistent across the years.

Figure L5. Were Reading Recovery students served in 1992-93 significantly less likely to need special assistance the following fall compared to groups served earlier?

Level of Service	Compared to (Year)	1992-93 cohort less likely to need:			
		Retention?	Special Ed?	Chapter 1?	Any Assistance?
Full Service	1990-1991	No	No	No	No
	1991-1992	No	No	No	No
Successful	1990-1991	No	No	No	No
	1991-1992	No	No	No	No
Not Successful	1990-1991	No	No	No	No
	1991-1992	No	No	No	No
Partial Service	1990-1991	Yes	No	No	No
	1991-1992	Yes	No	No	No
Total	1990-1991	No	No	No	No
	1991-1992	Yes	No	No	No

Note: Fisher's Exact Test was used to test significance, with a p. value of .05 or less set as the criteria. All shaded boxes show significant difference between the 1992-93 cohort and earlier cohorts. Yes is in favor of the 1992-93 cohort. No is in favor of the earlier cohort.

DISCUSSION

Special education placement of all RR groups in the year following RR participation has increased steadily since the program began. This mirrors an overall increase in the percentage of special education placement rates in WCPSS. The overall increase in placements is therefore not as much a concern as the fact that special education rates did not differ significantly between the RR and control groups. While some patterns of service did appear to be more positive for later RR groups, it is important to realize that the number of students involved also increased, which would make statistical significance easier to obtain in later years.

LONG-TERM EFFECTS ON READING ACHIEVEMENT

Q3. Do students who are successful in Reading Recovery stay at grade level in reading subsequently?

Only about one third of the 1990-91 and 1991-92 RR students showed grade-level proficiency in third grade on the EOG Reading test. Successful RR students did not perform significantly better than those who were less successful, and RR students overall did not perform better than comparison students (who did not participate in RR).

Similar to EOG results, 30% of 1990-91 RR successful students were able to score above the 50th percentile on the California Achievement Test (CAT). Only one fourth scored above the 50th percentile on the language subtest. A slightly higher percentage of RR students scored above the 50th percentile on both reading and language than the control group.

PROCEDURES AND LIMITATIONS

Some other RR sites have monitored subsequent student test scores for students who completed RR versus either a random band of students in the same school system or a comparison group. The Clay Observation Survey (developed by RR's founder) was sometimes used to monitor this growth, and nationally normed achievement tests were used in other studies. The Clay Observation Survey emphasizes oral reading ability, vocabulary, concepts of print, and simple writing skills, while norm referenced tests focus more on comprehension, vocabulary, and specific language skills. It would be ideal to know how the same students performed on both types of instruments. However, the testing time needed to complete this work is fairly prohibitive.

In WCPSS, two studies were done comparing third-grade performance of the RR students from 1990-91 with their control group. One study was conducted by program staff and a second by E&R staff. The control group was *very* well matched and similar to the RR group in the fall of 1990 because of the method of assignment. (In each school which offered RR services, the students were ranked in order of need and those with the greatest needs were then alternately assigned to either RR or to the control group.)

The RR staff examined the CAT Total Reading and Total Language percentile scores of *successful* RR students as they compared to a selected sample (1-to-1 matched with those students by first-grade pretest) of the control group. Third grade CAT scores were available for 37 of the 47 successful RR students and 33 of their 47 comparison students. They determined what percentage of students were able to score above the 50th percentile on each test. One limitation to the analysis was that only successful students were included. Another limitation is that the power of matching was diminished because when one of a pair was missing data, the other member of the pair was still kept in the analysis. Reading Recovery staff also looked at End-of-Grade (EOG) Reading state percentiles, but these will not be

reported here because they are less reliable to use for this type of analysis than the scale scores utilized in this report.

In the second study, the E&R Department compared the mean third-grade scores on the state's EOG Reading test for this same RR cohort (59 students) and their control group (68 students). In comparison with the CAT, the EOG test places greater emphasis on comprehension and higher order reading skills. We examined two types of EOG Reading scores: the developmental scale scores and the proficiency level scores. The developmental scale score is a normalized score with equal intervals; it is a tight scale with 100 points across grades 3-8. The proficiency level scores are based on teacher ratings during the test norming indicating student's level of mastery of essential grade-level skills. We examined the percentage of students scoring at Levels 3 or 4 (signifying mastery of grade-level reading skills and readiness for grade 4). Scores at Levels 1 and 2 indicate poor to inconsistent mastery of grade-level skills.

One additional analysis was completed in WCPSS comparing the 1991-92 RR and comparison students scoring at proficiency Levels 3 and 4 on the EOG Reading test.

One limitation for WCPSS and many other studies is that test scores on the same measures were not available at grades 1 and 3. Therefore, both WCPSS studies assume that a similar percentage of students would show scores "on grade level" on a norm-referenced test (CAT or EOG) as did students on the Clay Observation Survey used at grade 1.

RESULTS

One of the foundational longitudinal studies on the success of RR is an Ohio State University study which tracked performance of the Columbus, Ohio cohort who received RR tutoring in 1985-86 from grade 1 to grade 4. Clay Observation Survey performance was compared to the performance of students who began first grade at comparable achievement levels but who received regular Chapter 1 services (achievement comparison group) and to the performance of students who represented the range of achievement in the schools where RR was instituted (school comparison group). Secondary analysis of the Clay Observation Survey data by Wasick and Slavin (1993, Table 4 as cited by Hiebert), and then again by Heibert (1994), showed a positive effect of RR over regular Chapter 1 services on the Clay Observation Survey which decreased from +.78 of a standard deviation in grade 1 to +.25 in grade 3. At grade 4, the difference between RR students and the achievement comparison group was not statistically significant.

Some regional studies have shown a continued effect of RR in later years. Researchers at New York University compared 174 second-grade children who had been successful in RR to a grade-level random sample of 177 children. The RR children performed significantly better than the random sample on the Clay Observation Survey in text-reading skills at the end of second grade.

Using only successful RR students in WCPSS, RR staff found that 32.4% of the successful students and 27.2% of the comparison students were able to perform above the 50th percentile on the CAT Total Reading Test in the spring of grade 3. When the CAT Total Language percentiles were compared, it was found that 24.3% of the RR students and 12.2% of the control group performed above the 50th percentile. It was not reported whether these differences between RR and control students were statistically significant.

Evaluation and Research staff analyzed EOG scores for the two groups. On the state's EOG test, the RR students were slightly less likely (although not significantly so) than the control group students to score on grade level or above (32% vs. 34%, Fisher's Exact Test, p=.65). When comparing developmental scale scores, the RR group's mean was also slightly lower (but not significantly so) than that of the control group (135.5 vs. 136.2, p=.64). *Perhaps even more important, no significant differences were found when control group students were compared to only those RR students who received a full program or only to those students who were successful.*

Figure LR1. Third-Grade 1992-93 End-of-Grade Reading Scores*

	Number in Group	Number and Percent Scoring At Levels 3 or 4	Scale Score Means
1990-91 Reading Recovery Students			
Full Service	44	17 (39%)	136.4
Successful	36	13 (36%)	136.7
All Reading Recovery	59	19 (32%)	135.5
Control Students	68	23 (34%)	136.2

* Students' first third-grade EOG score was used; a few retainees' scores are from 1993-94.

The percentage of RR students scoring at each level was very similar to that of the control students. Approximately one quarter of each group scored at Level 1.

The 1991-92 RR and comparison groups scored similarly on the third-grade EOG Reading test. Just as the 1990-91 cohort comparison, one third of both groups scored at Levels 3 and 4, and two thirds fell in Levels 1 and 2. Specifically, 30% of comparison students, scored at Levels 3 and 4, while 28% of RR students scored at these levels. Within RR, 30% of those who received a full program and 33% of those who were successful scored at or above grade level (in Levels 3 and 4) on the EOG Reading. No significant differences were found between groups.

Thus, on both CAT and EOG, about one third of the RR students were able to show grade-level performance, and reading performance did not greatly exceed that of the control group.

Reading Recovery students appear to show some advantage over control students on the CAT, especially on the language section.

DISCUSSION

The fact that RR and comparison students from both 1990-91 and 1991-92 showed very similar performance on EOG at third grade is of great concern. The EOG is used as the State standard for "grade-level" performance at grades 3-8, and all those who do not score at Levels 3 and 4 will now be required to receive some focused intervention. Two thirds of *both* the RR and control groups would need further assistance by this criteria. Reading Recovery plus subsequent instruction was not enough to better prepare these students for the EOG at grade 3 than the control group students in the same schools. This supports the contention that it is difficult for any one intervention (especially a pull-out) to have a long-term impact unless classroom teachers effectively reinforce and build upon the principles of the intervention.

COST EFFECTIVENESS

Q4. Is Reading Recovery a cost effective way to help our students who have the most difficulty learning to read to succeed in school?

The average RR teacher serves seven students during a year, and, on average, three or four of those students read at a first-grade level by the end of the year. Annually, the cost per student for *all* students served in RR in WCPSS during 1993-94 was approximately $2,947.50 *beyond* the regular instructional program. The cost per *successful* student was approximately $6,000 beyond the regular instructional program. Current evaluation data suggests that by the end of third grade only about two of the students served by a RR teacher read at a third-grade level. Thus, the WCPSS has invested approximately $9,211 for each student who is a long-term success.

Since the 1990-91 and 1991-92 comparison groups of students who did not receive RR achieved a comparable success rate on standardized tests in third grade, and since RR expenditures in WCPSS do not seem to have been offset by significant savings from a reduction of need for special education, retention, or Chapter 1 assistance, the program does *not* appear to be cost effective at this time.

PROCEDURES AND LIMITATIONS

Several important questions must be asked in a discussion of cost effectiveness:

- Is it possible to meaningfully measure the outcomes of a program and attach a value to those outcomes?
- Is it possible to determine the approximate cost of implementing the program?
- Do alternative programs exist which have different costs and might produce comparable outcomes?

Incomplete answers to questions such as these impact any discussion of the cost effectiveness of RR. Among the problems that must be considered are the following:

- Reading Recovery is designed to bring the lowest achieving first-grade students in a school up to an average reading achievement level by the end of first grade, but "average" performance varies from school to school. The presence of a high percentage of high achieving students in a school may mean that it is more difficult to help low achievers reach an average level.
- "Lowest" is defined by RR to be the lowest 20% of first-grade students. If program services reach only the lowest 5-10%, then the success rate is likely to be lower than if services are provided to 15-20% of students.

- Various ways exist to measure the effectiveness of reading instruction. The RR program uses instruments that are individually administered and focus on oral reading accuracy, concepts of print, and basic writing skills. North Carolina's EOG reading achievement tests are group administered and focus on comprehension.

In attempting to assess the cost effectiveness of RR in WCPSS, start-up costs of training teachers were ignored, and only the long-term implementation cost of teacher salaries and benefits was considered. In addition, no attempt is made in this report to assess the use of time in each RR teacher's instructional day that lay outside the half-time commitment to RR. Thus, cost figures reported here could be considered an underestimate of the true total costs. Results discussed below are based upon average WCPSS teacher salary and benefits for 1993-94 of $41,265, or $20,632.50 for the half time spent on RR by each teacher in the program (serving an average of seven students).

RESULTS

Two nationally published discussions of RR's costs and benefits reached different conclusions. Philip C. Dyer, an elementary school principal writing in the winter 1992 issue of ERS Spectrum, concluded that the cost per student of RR tutoring was $2,063 and that school districts would derive long-term savings of $13,244 for each half-time teacher providing RR. Elfrieda Hiebert, a professor in the University of Michigan School of Education, reported in the December 1994 issue of Educational Researcher that the cost per successful student was $8,333 and that long-term savings cannot be estimated based upon available information.

Dyer's article made several assumptions that are not supported by WCPSS evaluation data. He assumed that each RR teacher serving eight children would prevent two first grade retentions, which implies a 25% retention rate for the RR target group if RR is not used. WCPSS data showed that the retention rate for students served in RR and for comparable students in a control group was only 5-6%. Dyer assumed that no students served by RR would be placed in special education programs, and that without RR, 12% of targeted students would be identified as Learning Disabled. WCPSS data showed that special education placement rates for both RR and the control group was approximately 12%.

Hiebert based her analysis upon the data available from three RR training sites which included student success rates and student-teacher ratios. She reached the following conclusions:

- In 1992, the average cost across several RR sites of successfully tutoring one first-grade student until the student reached an average first-grade reading achievement level was $3,488 (the cost per successful student). (This is based upon 86% of RR students successfully completing the program and one FTE teacher serving 11 students during the school year.)

- By the time they reached fourth grade, only 36% of RR students in a longitudinal study conducted in Ohio were still reading at an average level. If long-term success is the objective, the cost per successful student rises from $3,488 to $8,333.

Hiebert's approach of analyzing the cost for successful outcomes, when applied to WCPSS data and salary figures, yields the following conclusions:

- The cost per student for all students served in RR in WCPSS during 1993-94 was approximately $2,947.50 (over and above regular per pupil costs).
- Since approximately 50% of the students served from 1990 to 1994 were successful because they reached the average first-grade achievement level, the cost per successful student (short-term success) was approximately $6,000.
- The 1990-91 follow-up study in WCPSS indicated that 32% of all RR students scored at or above grade level in reading on the EOG at the end of third grade (2.24 of 7 students per RR teacher). Therefore, the cost for long-term success can be estimated at $9,211 per student, assuming that early RR assistance is the primary reason the students are successful at the end of third grade. *However, it must be noted that one third of the control group, who received no RR, also scored at or above grade level on the EOG in reading in third grade.*

Many factors beyond the control of RR staff affect whether or not a child is a successful reader at the end of third grade. For example, other WCPSS evaluation reports have discussed the possibility that second-grade teachers may not fully know how to build upon successful RR strategies, and that gains made in first grade may disappear in subsequent years. Nevertheless, if long-term success is the goal of any early childhood educational intervention, then the ultimate value and cost/benefit of RR must be judged by how students perform years after receiving RR services.

It is possible that somewhat higher success rates and lower costs per successful student could be obtained if the program served a larger percentage of the lowest readers in first grade. Currently only 4.4% of WCPSS students participate in the program. Serving a higher proportion of students would, presumably, direct services to students with problems less difficult to overcome and might result in a higher proportion of short-term and long-term success. The number of students served by each teacher might rise from six or seven to eight because the average student would require fewer tutoring sessions. The impact of broader coverage could be evaluated without undertaking major expenditures if existing RR staff were assigned to fewer schools and more students in each school were served. This would be our recommendation for 1995-96.

If the program is expanded, however, Figure C1 shows the estimated cost in 1993-94 dollars of several levels of expanded service. The costs shown assume that each teacher could tutor at least eight students during a school year, which is more than the six or seven students currently served, and that the program is implemented in all elementary schools.

Figure C1. Estimated Costs for Various Levels of Coverage

Level of Coverage	Students Served	Teachers Required	Cost
1993-94 Actual Coverage (4.4%)	292	41	$845,933
10% Coverage	670	84	$1,733,130
15% Coverage	1005	126	$2,599,695
20% Coverage	1340	168	$3,466,260

Advocates of RR assert that the cost of providing the program is, at least partially, offset by a reduction in the expenses associated with student retention, special education, and remedial services. Further research is needed before any cost savings can be calculated in WCPSS. Results the year after service showed some inconsistent benefits for RR over comparison groups (with the 1990-91 cohort less likely to need Chapter 1 and the 1991-92 less likely to be retained). However, no differences were evident two years after RR service (based on the 1990-91 group).

The lower need of RR students for Chapter 1 services in second grade could be considered a positive cost benefit since some Chapter 1 resources could be directed to additional students during that next year.

Even if the cost of RR is approximately $9,211 per long-term success, and no offsetting cost reductions can be found, the benefits obtained in helping a large number of beginning readers achieve success might make the investment appropriate if there was reason to believe that RR students would not learn to read without the help of the program. However, in WCPSS, about the same percentage of students in control groups scored at grade level on the third and fourth-grade EOG Reading test (28-34%) as did RR students (32%). This suggests that similar long-term success rates can be achieved by schools with or without the RR program and that the program cannot be considered cost effective at the present time.

SUMMARY OF TRENDS

Improving the skills of our lowest achievers continues to be one of the greatest challenges faced by our educational system. Early intervention programs like RR hold promise but are costly. Results for WCPSS' RR program thus far do not indicate it has been successful in fulfilling its goals to:

- Bring the lowest first graders to grade level in reading;
- Keep RR students at grade level (with less need for further help than would be true otherwise); and
- Reduce the need for special education services or retention.

Actual WCPSS results have shown that:

- About 50% of those served--two thirds of those receiving a full program--have been brought to grade level in first grade. All but the 1993-94 results are fairly consistent with the range of success rates in NC and other national sites. There has been a decline in WCPSS RR success rates each year.
- Only about one third of 1990-91 and 1991-92 RR students--about the same as the comparison groups not served by RR--were at grade level on the CAT and EOG in third grade. Some longitudinal studies at other U.S. sites have found more promising results.
- Reading Recovery students from 1990-91 were not less likely than the control group to be retained, placed in special education, or served by Chapter 1 two years later. About half of the RR students have needed some form of additional assistance.

While some WCPSS students have clearly benefited from RR, the program overall has not been as successful to date as hoped. Ways to strengthen the program's short- and long-term success should be implemented in such a way that their success can be studied in a systematic way. If feasible, full coverage should be provided in at least a few RR schools and compared to the success of at least one other promising early intervention approach in a few other schools.

REFERENCES

Donley, J., & Baenen, N. (1993). Evaluation report: Reading Recovery 1990-92 (Report No. 93.09). Raleigh, North Carolina: Wake County Public School System, Department of Evaluation and Research.

Dyer, P. C. (1992). Reading Recovery: A cost-effectiveness and educational-outcomes analysis. ERS Spectrum, 10(1), 10-19.

Hiebert, E. H. (1994, December). Reading Recovery in the United States: What difference does it make to an age cohort? Educational Researcher, 23 (9), pp.13-25.

Hundley, S. (1994, April). Research findings follow-up of the 1990-91 cohort. (Available from Sue Hundley, Wake County Public School System, Reading Recovery Office.)

Jones, N.K. (1994), Reading Recovery in North Carolina: A three-year report 1990-93. Wilmington: University of North Carolina at Wilmington, School of Education.

Klein, S. (1993). Are there better ways to identify and share the best within and among federal education programs? Evaluation and Program Planning, 16(3), 227-39.

Pinnell, G. S., Lyons, C., & DeFord, D. (1993). Reading Recovery executive summary 1984-1993. Columbus, OH: Ohio State University, Reading Recovery Program.

Pinnell, G. S., Lyons, C., & DeFord, D. (1994). Reading Recovery executive summary 1984-1994. Columbus, OH: Ohio State University, Reading Recovery Program.

Pinnell, G. S., Lyons, C. A., DeFord, D. E., Bryk, A. S., & Seltzer, M. (1994). Comparing instructional models for the literacy education of high-risk first graders. Reading Research Quarterly, 29(1), 9-39.

Pollock, J.S. (1991). Reading Recovery program 1990-91: Final Evaluation Report. Elementary and Secondary Education Act--Chapter 1. Ohio: Columbus Public Schools, Department of Program Evaluation.

Pollock, J.S. (1993). Reading Recovery program 1991-92: Elementary and Secondary Education Act--Chapter 1. Final Evaluation Report. Ohio: Columbus Public Schools, Department of Program Evaluation.

Walberg, H., & Niemiec, R. (1993). Validating exemplary programs: Methods and criteria. Educational Evaluation and Policy Analysis, 15(4), 429-36.

Wasik, B.A., & Slavin, R.E. (1993). Preventing reading failure with one-to-one tutoring: A review of five programs. Reading Research Quarterly, 28, 178-200.

ATTACHMENT 1.

Retention Rates Before and After Reading Recovery for Grade 1 as of August

Retention Rates Are First Shown for Each School the Year Before Reading Recovery Was Implemented

WCPSS Schools	1989-90		1990-91		1991-92		1992-93		Comparison before RR and after 1st Year *	Comparison before RR and after 2nd Year *
	#	%	#	%	#	%	#	%		
Adams					0	0.00	0	0.00	NA	
Aversboro					0	0.00	0	0.00	NA	
Baucom	2	2.19	1	0.81	0	0.00			Y	Y
Brentwood			1	1.01	2	2.02	5	4.81	N	N
Bugg	1	1.38	1	1.25	1	1.61			Y	N
Carver			4	2.82	0	0.00	0	0.00	Y	Y
Cary					0	0.00	0	0.00	NA	
Conn			0	0.00	1	0.99	1	0.83	N	N
Creech Road/Garner	1	0.86	1	1.04	1	1.15			N	N
Douglas	3	2.52	2	2.27	1	0.92			Y	Y
Farmington Woods					2	1.58	1	0.70	Y	
Fox Road					6	3.95	0	0.00	Y	
Fuller	2	2.56	0	0.00	0	0.00			Y	Y
Green					0	0.00	1	1.70	N	
Hunter					1	0.95	4	3.81	N	
Joyner					0	0.00	0	0.00	NA	
Kingswood**	4	8.16	0	0.00	0	0.00	0	0.00	Y	
Knightdale	15	6.97	3	1.62	7	3.68			Y	Y
Lacy			6	4.96	3	2.73	6	4.88	Y	Y
Lincoln Heights			0	0.00	1	0.81	0	0.00	N	NA
Lynn Road					1	0.81	0	0.00	Y	
Millbrook					1	0.75	0	0.00	Y	
Poe	0	0.00	0	0.00	0	0.00			NA	NA
Powell			3	4.23	1	1.37	0	0.00	Y	Y
Rand Road			2	1.96	1	1.03	0	0.00	Y	Y
Rolesville	2	2.15	3	3.53	2	1.75			N	Y
Smith					1	1.04	1	0.98	Y	
Stough	2	1.58	1	1.03	1	1.10			Y	Y
Vance			0	0.00	1	1.09	0	0.00	N	NA
Vandora Springs			2	2.25	3	3.09	2	2.33	N	N
Wake Forest					9	4.55	11	6.92	N	
Washington					1	1.18	0	0.00	Y	
Wilburn					0	0.00	0	0.00	NA	
Wiley	2	3.77	0	0.00	1	1.30			Y	Y
Willow Springs			3	5.17	1	1.47	4	4.82	Y	Y
Zebulon					15	13.27	4	3.23	Y	
Districtwide Totals	173	3.25	103	1.90	96	1.67	95	1.55		

Note: Retention rates for 1993-94 are not available because the report, previously done by Long Range Planning, is not available.

* For the two comparison columns; a "Y" means retention rates decreased across years, "N" shows that retention rates increased across years, and "NA" refers to changes that were not applicable because the rates were both at 0.

** Kingswood Elementary did not participate in the Reading Recovery program in 1991-92, but did in 1990-91 and 1992-93.

Attachment 2. Reading Recovery Cohorts

Proportion Receiving Subsequent Assistance in the Next School Year

Type of Reading Recovery Student	Cohort		
	1990-91	1991-92	1992-93
Proportion Retained			
Full Service	1/51 (2%)	4/101 (4%)	3/155 (2%)
Successful	0/39 (0%)	2/77 (3%)	0/113 (0%)
Not successful	1/12 (8%)	2/24 (8%)	3/42 (7%)
Partial Service	3/21 (14%)	7/40 (18%)	1/73 (1%)
Total	4/72 (6%)	11/141 (8%)	4/228 (2%)
Proportion Served by Special Education			
Full Service	3/51 (6%)	22/101 (22%)	43/155 (28%)
Successful	2/39 (5%)	13/77 (17%)	23/113 (20%)
Not successful	1/12 (8%)	9/24 (38%)	20/42 (48%)
Partial Service	6/21 (29%)	10/40 (25%)	35/73 (48%)
Total	9/72 (13%)	32/141 (23%)	78/228 (34%)
Proportion Served by Chapter 1			
Full Service	17/51 (33%)	32/101 (32%)	42/155 (27%)
Successful	10/39 (26%)	19/77 (25%)	28/113 (25%)
Not successful	7/12 (58%)	13/24 (54%)	14/42 (33%)
Partial Service	6/21 (29%)	17/40 (43%)	30/73 (41%)
Total	23/72 (32%)	49/141 (35%)	72/228 (32%)
Proportion Assisted in Any of These Ways			
Full Service	19/51 (37%)	51/101 (51%)	76/155 (49%)
Successful	11/39 (28%)	30/77 (39%)	48/113 (42%)
Not successful	8/12 (67%)	21/24 (88%)	28/42 (67%)
Partial Service	12/21 (57%)	31/40 (78%)	55/73 (75%)
Total	31/72 (43%)	82/141 (58%)	131/228 (57%)

Proportion Receiving Subsequent Assistance in the Next School Year
Special Education Service for Learning Disabled (LD) Only

Type of Reading Recovery Student	Cohort		
	1990-91	1991-92	1992-93
Proportion Retained			
Full Service	1/51 (2%)	4/101 (4%)	3/155 (2%)
Successful	0/39 (0%)	2/77 (3%)	0/113 (0%)
Not successful	1/12 (8%)	2/24 (8%)	3/42 (7%)
Partial Service	3/21 (14%)	7/40 (18%)	1/73 (1%)
Total	4/72 (6%)	11/141 (8%)	4/228 (2%)
Proportion Served by Special Education (LD only)			
Full Service	1/51 (2%)	13/101 (13%)	15/155 (10%)
Successful	1/39 (17%)	7/77 (9%)	4/113 (4%)
Not successful	0/12 (0%)	6/24 (25%)	11/42 (26%)
Partial Service	2/21 (10%)	9/40 (23%)	21/73 (29%)
Total	3/72 (4%)	22/141 (16%)	36/228 (16%)
Proportion Served by Chapter 1			
Full Service	17/51 (33%)	32/101 (32%)	42/155 (27%)
Successful	10/39 (26%)	19/77 (25%)	28/113 (25%)
Not successful	7/12 (58%)	13/24 (54%)	14/42 (33%)
Partial Service	6/21 (29%)	17/40 (43%)	30/73 (41%)
Total	23/72 (32%)	49/141 (35%)	72/228 (32%)
Proportion Assisted in Any of These Ways			
Full Service	19/51 (37%)	45/101 (45%)	55/155 (35%)
Successful	11/39 (28%)	26/77 (34%)	32/113 (28%)
Not successful	8/12 (67%)	19/24 (79%)	23/42 (55%)
Partial Service	11/21 (52%)	30/40 (75%)	48/73 (66%)
Total	30/72 (42%)	75/141 (53%)	103/228 (45%)

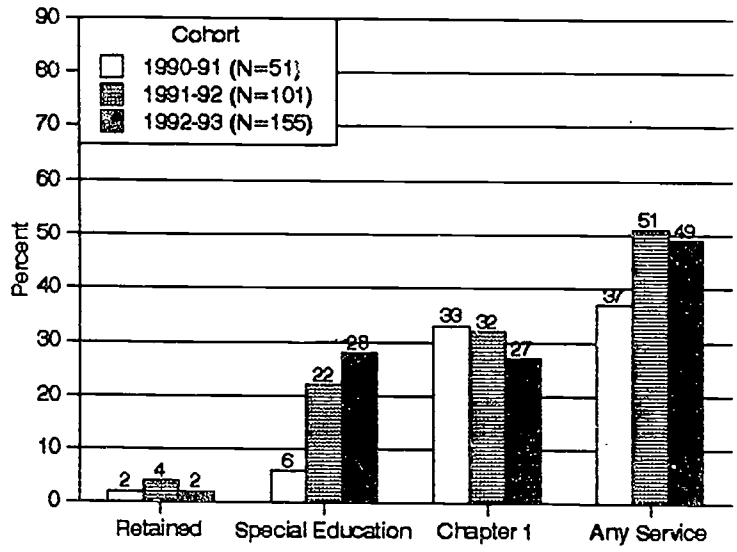
APPENDIX A: SUBGROUP RESULTS ON FURTHER ASSISTANCE NEEDS

Full versus Partial Service

RR students who received full services were significantly less likely than those who received partial services to need some form of subsequent assistance for two of the three groups studied (the 1991-92 and 1992-93 groups but not the 1990-91). When examined by types of assistance, the pattern was inconsistent, with only four of nine comparisons favoring students who received full services. The pattern remained similar when only special education placement into the Learning Disabled category was considered.

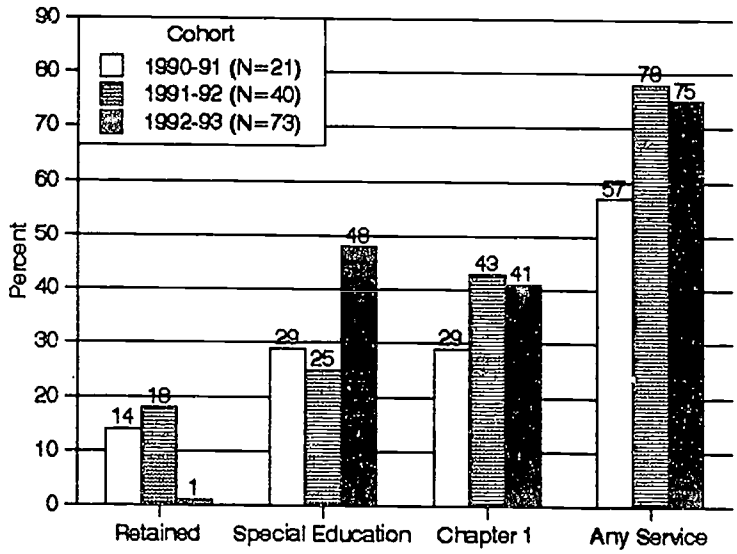
Patterns varied slightly across years for students who received full and partial service through RR as shown in the figures which follow.

Reading Recovery Full Service Students: One Year Follow-Up



Among *full-service* students, the only significant difference across years was found in comparing the rates of special education placement, with overall rates significantly *higher* for the 1991-92 and 1992-93 cohort than the 1990-91 cohort. When only Learning Disabled was considered, this significance was lost. While the Chapter 1 placement declined, differences were not significant. (See Table 1.)

Reading Recovery Partial-Service Students: One Year Follow-Up



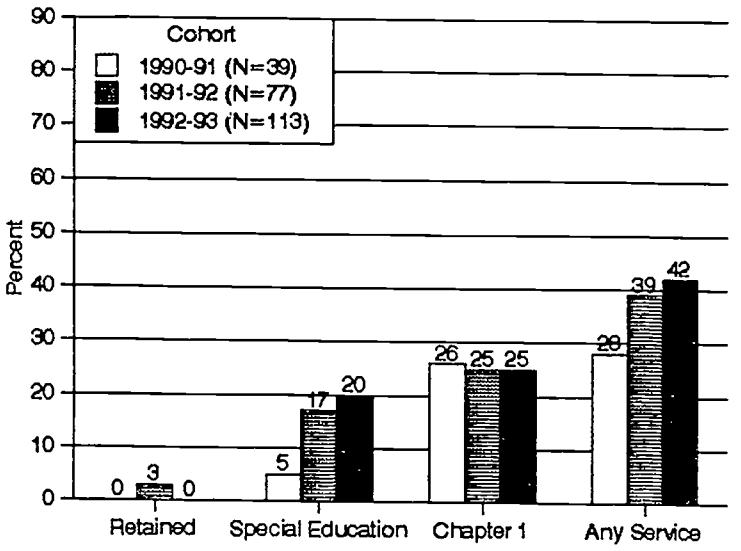
Among partial-service students, the 1992-93 cohort, as with full-service students, had a significantly *larger* proportion placed in *special education* programs than the two previous cohorts. Placements in the Learning Disabled category did not differ significantly across years. The 1992-93 cohort did have a significantly smaller proportion of students *retained*.

Successful versus Not Successful Students

Within full-service students, patterns for students who were successful in RR were examined as well as those who were not.

Successful students were significantly less likely to need some type of subsequent assistance than unsuccessful ones in all three cohorts. When examined by type of assistance, successful students were significantly less likely to be retained in grade or placed in special education or Chapter 1 in only four of nine comparisons. Patterns were not consistent across type of service. Patterns for 1991-92 and 1992-93 *were* more positive than in 1990-91. (See Table 2.)

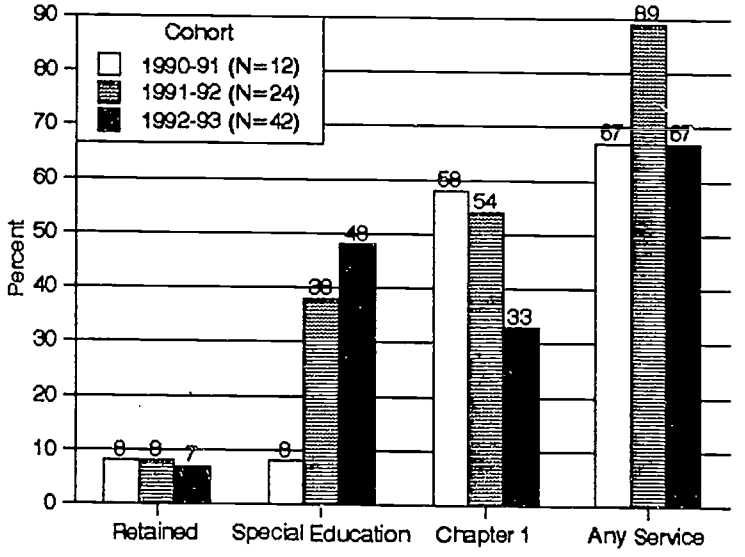
Reading Recovery Successful Students: One Year Follow-Up



Note: None of the successful students were retained from the 1990-91 or the 1992-93 cohort.

For successful students, the only significant difference was in the comparison of special education placements. The 1992-93 cohort was *more* likely to be placed in special education than the 1990-91 cohort. However, they were *not* more likely to be placed in special education for Learning Disabilities.

Reading Recovery Not Successful Students: One-Year Follow-Up



Among students who received full services but were not successful, there were *no significant changes* when the *1992-93* cohort was compared with the 1991-92 cohort, *although* the rates for special education program placement were *higher for both these cohorts than for the 1990-91 cohort*. The difference in the percentage of students needing any service in 1992-93 (67%) and 1991-92 (88%) approached significance ($p=.055$). When only the Learning Disabled category

was considered, special education placements in 1992-93 were similar to 1991-92 (but still higher than 1990-91), but the overall percentage of students receiving any service was lower in 1992-93 than 1991-92 (55% versus 79%, p=.04). For further information see Attachment 2.

Table 1. Were Reading Recovery students who received full services less likely to need subsequent assistance than those with partial service?

Year	Full service students less likely to need:			
	Retention?	Special Ed?	Chapter 1?	Any Assistance?
1990-91	No	Yes (0.02)	No	No
1991-92	Yes (0.01)	No	No	Yes (0.004)
1992-93	No	Yes (0.004)	Yes (0.05)	Yes (<0.001)

Note: p values are based on the Fisher's Exact Test

Table 2. For those who received full services, were successful students less likely to need subsequent assistance than not successful ones?

Year	Successfully-discontinued students less likely to need:			
	Retention?	Special Ed?	Chapter 1?	Any Assistance?
1990-91	No	No	No	Yes (0.04)
1991-92	No	Yes (0.05)	Yes (0.01)	Yes (<0.001)
1992-93	Yes (0.02)	Yes (0.001)	No	Yes (0.01)

EVALUATION REPORT: WCPSS READING RECOVERY 1990-94

Authors
Because of the scope of this project, a number of staff contributed.

Nancy Baenen
Director of Evaluation

Alissa Bernholc
Evaluation Consultant

Chuck Dulaney
Evaluation Specialist

Karen Banks
Assistant Superintendent of Evaluation and Research

Margaret Willoughby
Evaluation Technician

April 1995
E&R Report No. 95.09A

Wake County Public School System
Department of Evaluation and Research