

DOCUMENT RESUME

ED 430 067

UD 032 910

AUTHOR Huggins, Regina
 TITLE Longitudinal Study of the Reading Recovery Program, 1994-1998.
 INSTITUTION Detroit Public Schools, MI. Office of Research, Evaluation, and Assessment.
 PUB DATE 1999-04-00
 NOTE 27p.
 PUB TYPE Numerical/Quantitative Data (110) -- Reports - Evaluative (142)
 EDRS PRICE MF01/PC02 Plus Postage.
 DESCRIPTORS *Achievement Gains; *Early Intervention; *Grade 1; Longitudinal Studies; Primary Education; Program Effectiveness; Program Evaluation; Reading Achievement; Reading Tests; Standardized Tests; Tables (Data); Urban Schools
 IDENTIFIERS *Detroit Public Schools MI; *Reading Recovery Projects

ABSTRACT

Reading Recovery is an early intervention program for first grade designed to reduce reading failure. It is based on the assumption that intensive, high-quality help during the early years of schooling is the most productive investment of resources. Individually tailored intervention procedures are implemented for children identified by classroom teachers as achieving below-grade level. A longitudinal study was designed to assess the success of this program in the Detroit Public Schools (Michigan) over the 5 years from 1993-94 to 1997-98. A total of 135 students were studied, 75 in the experimental group participating in Reading Recovery and 60 in a comparison group who were approximately at the same reading achievement level at the start of the study. Data from the California Achievement Test, the Metropolitan Achievement Test, and the Michigan Education Assessment Program were used to measure student achievement. The Reading Recovery group showed improvement throughout the study--an improvement sustained over the 5 years. However, the Reading Recovery group consistently scored below the achievement levels of the comparison group. It is recommended that the program, if continued, be monitored closely for documentation of academic achievement and effectiveness. An appendix contains information about assessment tasks administered in Reading Recovery. (Contains 20 tables.) (SLD)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

LONGITUDINAL STUDY OF THE READING RECOVERY PROGRAM 1994 - 1998

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL HAS
BEEN GRANTED BY

J. Moore
D. P. S.

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

1

**Regina Huggins, Ph.D.,
Evaluator**

**OFFICE OF RESEARCH, EVALUATION AND ASSESSMENT
DIVISION OF EDUCATIONAL SERVICES
DETROIT PUBLIC SCHOOLS
APRIL, 1999**

TABLE OF CONTENTS

Executive Summary	i
Program Facts	iv
Program Description/Evaluation Methodology	1
Findings	3
Conclusion	15
Recommendation	16
Appendix 1a – 1c	17

LONGITUDINAL STUDY OF THE READING RECOVERY PROGRAM

Regina Huggins, Evaluator

Executive Summary

Purpose and Features of the Program

Reading Recovery is an early intervention program for first grade students designed to reduce reading failure. It is based on the assumption that intensive, high-quality help during the early years of schooling is the most productive investment of resources. In Reading Recovery, individually administered diagnostic procedures are used to identify children in need of special help. These children are identified through classroom teacher assessment. Intervention procedures are tailored individually to help a child working below grade level become a successful reader. Teachers in training are trained by lead teachers in the program. These lead teachers are experienced teachers who received a year of full-time graduate credit *to be certified to conduct training classes for Reading Recovery teachers.

Methodology

The Longitudinal Study of the Reading Recovery Program was designed to assess the success of the program over a five-year time frame (1993-94 to 1997-98). This longitudinal study involved observing the Reading Recovery group independently, as well as, a comparison analysis with participants from the Reading Recovery group and the control group. A total of 135 students were used in the study. The Reading Recovery staff in the Office of Communication Arts supplied the names for the study. Seventy-five first grade students were in the experimental group and were identified by kindergarten and/or first grade teachers as being the lowest children in the class. This list was reinforced by the Reading Recovery staff who administered their own *departmental assessment to verify the bottom 20-30% needing Reading Recovery tutoring. These students were enrolled in the program for one-full year (1993-1994) and were taught by teachers trained in Reading Recovery during the 1992-1993 school year. The remaining sixty students were in the control group. These students were randomly selected during the spring of 1994 from classrooms that had students receiving Reading Recovery. They were students with no exposure to Reading Recovery tutoring, and were identified through classroom teachers as performing slightly above the children selected for Reading Recovery to the top of the class. Data from the California Achievement Test (*CAT-E*) administered during the spring of 1993 (kindergarten), and the spring of 1994 (first grade); the Metropolitan Achievement Test (*MAT7*)

* *Clay's Assessment Tasks* - An assessment tool administered by the Reading Recovery teachers to the lowest 20 - 30% of the students as identified by Kindergarten and/or First Grade teachers, (Appendix 1a - 1c).

administered in the spring of 1995, 1996, 1997 and 1998; and the Grade 4 Michigan Education Assessment Program (*MEAP*) administered in the spring of 1997 were collected. The evaluator of the Reading Recovery Program was responsible for collecting and analyzing all data.

This evaluation involved a variety of statistical tests designed to assess the value of the Reading Recovery Program. A chi-square test of significance, frequency distribution analysis, as well as, sample statistic t-test were completed to determine if a significant difference existed between the Reading Recovery group and the control group over a five-year time frame.

Two criteria were used to measure the success of the Reading Recovery Program: First, students who were enrolled in the program for one-full year will maintain their level of reading ability over a four-year period (2nd through 5th grades). Second, the students who participated in Reading Recovery will have significantly higher reading scores than a comparison group of students who did not participate in Reading Recovery.

Findings

The Reading Recovery group showed reading improvement throughout the study when compared to the 1994-95 MAT7 results. Using frequency distributions of MAT7 data and 1994-95 as the base year, the findings indicate:

- For 1995-96, the Reading Recovery group had 54.5% of its students in the average range, an increase of 16.2% over 1994-95, while 3.0% scored in the above average range, a decrease of 2.9%.
- For 1996-97, the Reading Recovery group had 46.6% of its students in the average range, an increase of 8.3% over 1994-95, while 5.0% scored in the above average range, a decrease of 0.9%.
- For 1997-98, the Reading Recovery group had 54.3% of its students in the average range, an increase of 16.0% over 1994-95, and 5.3% scored in the above average range, a decrease of 0.6%.

Both groups, the Reading Recovery and control group, performed around the same level of achievement on both the spring 1993 CAT-E, as well as, the spring 1994 CAT-E. For the 1993 CAT-E assessment all students involved in the study were in kindergarten with no exposure to Reading Recovery. For the 1994 CAT-E, all students were in first grade. The experimental group had exposure to Reading Recovery, and the control group had no exposure to Reading Recovery. The difference in performance occurred on the MAT7 from spring 1995 to spring 1998 and on the 1996-97 Grade 4 MEAP. On these assessments, when first grade students who participated in Reading Recovery were compared to first grade students who had no exposure to Reading

Recovery, the findings indicated that a significant difference existed in favor of the control group.

Upon examining the MAT7 stanine scores and related categories of achievement, a significant difference existed between the two groups for each year beginning spring 1995 through spring 1997, in favor of the control group. For the spring 1998 MAT7, the percent of students in the average and above average categories was higher for the control group although the difference was not significant. The mean scale score results showed that the control group outperformed the Reading Recovery group in three of the four years of the study. Only in the spring of 1998 did the Reading Recovery group have a higher mean scale score.

Based on the findings of this longitudinal study over a five-year time frame (spring 1994 to spring 1998), Reading Recovery was successful in having its participants maintain their level of achievement. However, the Reading Recovery group consistently scored below the achievement levels of the control group. Often, these differences were significant.

Recommendation

Based upon this longitudinal study, the following recommendation is listed:

If the Reading Recovery Program is continued, it must be closely monitored for documentation of academic achievement and effectiveness.

PROGRAM FACTS

Name of Program	Reading Recovery
Funding Year	1993 - 1998
Purpose of Program	The goal of Reading Recovery is to accelerate learning. It is designed to move children in a short time from the bottom of their class to the average where they can profit from regular classroom instruction. Children are expected to develop a self-extending system using a variety of strategies to read increasingly difficult books and to independently write their own messages.
Program Description	Reading Recovery is an early intervention program designed to assist first grade children in reading acquisition. Grade 1 teachers observe students and rank them according to their reading performance. A list of the lowest 20-30 percent of first graders is formulated to determine their assessment needs. Reading Recovery provides one-on-one tutoring, five days per week, 30 minutes a day, by a specially trained teacher. The 30-minute daily lessons consist of a variety of reading and writing experiences designed to help children develop their own effective strategies for literacy acquisition. Reading Recovery lasts an average of 12-20 weeks and is supplemental to classroom instruction.
Funding Source	<i>1993 - 1996</i> Chapter 1 31A <i>ACE (Accelerating Change in Education)</i> <i>1997 - 1998</i> Title I - <i>Buy-In</i> Title I - <i>carry-over</i> 31A Board Funds Eisenhower Grant Funds
Number and Level of Participants	<i>*Supplied by the <u>Office of Communication Arts</u></i> First Graders 1993-94 = 116 1994-95 = 172 1995-96 = 254 1996-97 = 266 1997-98 = 396

Number and Level of Schools in Program	Thirty (30) Elementary Schools
Name of Schools	Beard, Bethune, Carstens, Clark, Clinton, Cooke, Cooper, Grayling, Guyton, Hanstein, A. L. Holmes, Houghten, J. R. King, Law, Loving, MAAT, MacDowell, Malcolm X, Marsh, J. Marshall, Monnier, Pasteur, F. Richards, Scripps, Stewart, Vandenberg, Vernor, White, Winterhalter, and Woodward.
Staffing Pattern	<i>*Supplied by the <u>Office of Communication Arts</u></i> 1993-94 = 22 teachers 1994-95 = 56 teachers 1995-96 = 67 teachers 1996-97 = 81 teachers 1997-98 = 74 teachers
Instructional Time	Half-hour daily for an average of 12-20 weeks
Equipment/Materials	Variety of children reading books, Clinical Pack, Diagnostic Assessment Book, Course Book, Reading Recovery level books (over 450), and teaching supplies.
First Year Funded	1992-1993

Longitudinal Study of the Reading Recovery Program 1994 - 1998

Program Description

Reading Recovery is designed as an early assessment, intervention, and observation program involving children in the first grade who are having difficulty learning to read and write. Classroom teachers identify students as the lowest in their class in reading acquisitions. Students are observed and ranked according to their performance. A list of the lowest 20 - 30 percent of first graders is formulated to determine the assessment tasks. Reading Recovery provides one-to-one tutoring, five days per week, 30-minutes a day, by a specially trained teacher.

Reading Recovery daily lessons consist of a variety of reading and writing experiences that are designed to help children develop their own effective strategies for literacy acquisition. Reading Recovery is supplemental to classroom instruction and lasts an average of 12-20 weeks. During the lesson the child is consistently engaged in holistic reading and writing tasks. Each lesson includes reading mini books, composing and writing a story. Every day the child is introduced to a new book, which he/she will be expected to read without help the next day. Strategies for hearing sounds in words and for monitoring and checking their own reading is developed through writing. Students are expected at the end of the series of lessons to have developed a variety of strategies designed to increase reading and to independently write their own messages.

Reading Recovery uses a trainer of trainer's model. Experienced teachers are provided with professional development in a yearly curriculum that integrates theory and practice and is characterized by intensive interaction with colleagues. Teachers-in-training conduct lessons behind a one-way glass and are observed and given feedback by their colleagues. The lead teacher visits teachers at their sites and assists with professional development. The professional development has three main elements: (1) extensive training program involving practice/observation/discussion; (2) coaching by teacher leaders; and (3) ongoing professional development-demonstrations/observation/and critique participation.

Methodology

The following questions were utilized throughout the longitudinal study:

- What effect did Reading Recovery have over a period of time in standardized testing?
- Did the students who originally entered the program maintain progress in the second, third, fourth, and fifth grades?
- How did the Reading Recovery students compare with a control group of students?

Methodology (continued)

Two criteria will be used to measure the success of the Reading Recovery program: First, students who were enrolled in the program for one-full year will maintain their level of reading ability over a four year period (*2nd through 5th grades*). Second, the students who participated in Reading Recovery will have significantly higher reading scores than a comparison group of students who did not participate in Reading Recovery.

To assess the first criteria, the Metropolitan Achievement Test (*MAT7*) for spring of 1995, 1996, 1997 and 1998 data were collected, analyzed, and summarized for the Reading Recovery group. Individual scale scores were converted to stanine scores for the purpose of reporting below average, average, and above average results from *MAT7*. Furthermore, frequency distributions were prepared to depict the numbers and percents of students scoring below average, average and above average data for the school years 1994-95 to 1997-98. The base year, 1994-95, was used as a guide for determining if the Reading Recovery group made improvements in their reading achievement as they advanced through the grades.

To assess the second criteria, the evaluation design included a comparison analysis for the Reading Recovery with a control group. The students' scores for the three test instruments administered during the 1993 to 1998 school year time period were analyzed.

- * *The California Achievement Test-Form E (CAT-E)* was administered during the spring of 1993 (kindergarten), and the spring of 1994 (first grade). The Total Reading scale scores were summarized for each group and a t-test of significance was conducted.
- * *The Metropolitan Achievement Test (MAT7)* was administered in the spring of 1995 to the spring of 1998. An independent samples test was conducted for each year, comparing the mean scale scores of each group. Stanine scores were categorized into "below average", "average", and "above average" categories and compared.
- * The Grade 4 *Michigan Education Assessment Program (MEAP)* was administered in the spring of 1997. A frequency distribution of the number and percent of students in each category of achievement on the MEAP reading test for each group was prepared. A Chi-square test of significance was conducted.

A total of 135 first grade students were observed in the longitudinal study. Seventy-five first grade students were in the experimental group and were identified by kindergarten and/or first grade teachers as being the lowest children in the class. This list was reinforced by the Reading Recovery staff who administered Clay's Six Assessment Tasks* observation survey to verify the bottom 20 – 30% needing Reading Recovery tutoring. These students were enrolled in the Reading Recovery program for one-full year (1993-1994) and were taught by teachers trained in Reading Recovery during the 1992-1993 school year. The remaining sixty students were in the control group. These students were randomly selected during the spring of 1994 from classrooms that had students receiving Reading Recovery. They were students with no exposure to Reading Recovery tutoring, and were identified through classroom teachers as

* Clay's Six Assessment Tasks – An observation survey of early literacy achievement administered by the Reading Recovery teachers to the lowest 20 – 30% of the students as identified by Kindergarten and/or First Grade teachers, (Appendix 1a – 1c).

performing slightly above the children selected for Reading Recovery to the top of the class. These printout sheets listing participants in the experimental and the control group were obtained from the Reading Recovery staff in the Office of Communication Arts.

Findings

The stanine scale is a convenient and useful basis for score interpretation. It derives its meaning from two words, "standard nine", which express its basic meaning. Stanine scores are normally distributed scores that range from 1 to 9. It is generally considered that stanine scores in the 1-3 range are "below average"; scores in the 4-6 range are "average"; and scores in the 7-9 range are "above average".

The MAT7 scale scores were converted to stanine scores. A year-by-year analysis of the MAT7 stanine scores follow.

Table 1 shows the 1994-95 Grade 2 test data from MAT7. Sixty-eight (90.7%) of the seventy-five Reading Recovery students had recorded MAT7 test scores. The group's mean score was 3.66, a *below average stanine score*.

Table 1
MAT7 Reading Recovery Reading Stanine Scores
(Grade 2)
1994-1995

Category	Number of Participants	Mean	Standard Deviation
Comprehension	68	3.74	1.39
Vocabulary	68	3.99	2.33
Total Reading	68	3.66	1.72

Table 2 shows the 1995-96 Grade 3 test data from MAT7. Sixty-six (88.0%) of the seventy-five Reading Recovery students had recorded MAT7 test scores. The group's mean score was 3.70, a *below average stanine score*.

Table 2
MAT7 Reading Recovery Reading Stanine Scores
(Grade 3)
1995-1996

Category	Number of Participants	Mean	Standard Deviation
Comprehension	66	3.79	1.43
Vocabulary	66	3.78	1.99
Total Reading	66	3.70	1.53

Table 3 shows the 1996-97 Grade 4 test data from MAT7. Sixty (80.0%) of the seventy-five Reading Recovery students had recorded MAT7 test scores. The group's mean score was 3.38, *a below average stanine score.*

Table 3
MAT7 Reading Recovery Reading Stanine Scores
(Grade 4)
1996-1997

Category	Number of Participants	Mean	Standard Deviation
Comprehension	60	3.57	1.52
Vocabulary	60	3.58	2.00
Total Reading	60	3.38	1.61

Table 4 shows the 1997-98 Grade 5 test data from MAT7. Fifty-seven (76.0%) of the seventy-five Reading Recovery students had recorded MAT7 test scores. The group's mean score was 4.05, *at the lower end of the average stanine score range of achievement.*

Table 4
MAT7 Reading Recovery Reading Stanine Scores
(Grade 5)
1997-1998

Category	Number of Participants	Mean	Standard Deviation
Comprehension	57	4.02	1.56
Vocabulary	57	4.32	2.18
Total Reading	57	4.05	1.77

Historical perspectives of the mean MAT7 stanine scores for comprehension, vocabulary and total reading are displayed in Table 5, Table 6, and Table 7 respectively. In each case, the mean was below 4.0 for the 1994-95, 1995-96 and 1996-97 test administration. For 1997-98, the mean score was slightly above 4.0.

BEST COPY AVAILABLE

Table 5
MAT 7 Comprehension Mean Stanine Scores for the
Reading Recovery Group

Year / Grade	Reading Recovery		
	Number	Mean	Standard Deviation
94-95 / 2	68	3.74	1.39
95-96 / 3	66	3.79	1.43
96-97 / 4	60	3.57	1.52
97-98 / 5	57	4.02	1.56

Table 6
MAT 7 Vocabulary Mean Stanine Scores for the
Reading Recovery Group

Year / Grade	Reading Recovery		
	Number	Mean	Standard Deviation
94-95 / 2	68	3.99	2.33
95-96 / 3	66	3.78	1.99
96-97 / 4	60	3.58	2.00
97-98 / 5	57	4.32	2.18

Table 7
MAT 7 Total Reading Mean Stanine Scores for the
Reading Recovery Group

Year / Grade	Reading Recovery		
	Number	Mean	Standard Deviation
94-95 / 2	68	3.66	1.72
95-96 / 3	66	3.70	1.53
96-97 / 4	60	3.38	1.61
97-98 / 5	57	4.05	1.77

The frequency distribution of the Reading Recovery Total Reading stanine scores for 1994-95 to 1997-98 MAT7 is displayed in Table 8. The category of achievement represented by the range of stanine scores is summarized in Table 9.

Table 8

**Frequency Distribution of MAT7
Total Reading Stanine Scores for
Reading Recovery Group**

Stanine	1994-95		1995-96		1996-97		1997-98	
	Number of Participants	%						
1	4	5.9	8	12.1	9	15.0	3	5.3
2	13	19.1	3	4.5	9	15.0	12	21.1
3	21	30.9	17	25.8	11	18.3	8	14.0
4	11	16.2	22	33.3	21	35.0	6	10.5
5	10	14.7	8	12.1	5	8.3	17	29.8
6	5	7.4	6	9.1	2	3.3	8	14.0
7	1	1.5	1	1.5	2	3.3	2	3.5
8	2	2.9	1	1.5	1	1.7		
9	1	1.5					1	1.8
Total	68	100.1	66	99.9*	60	99.9*	57	100.0

*Total does not equal 100 due to rounding.

Table 9

**Summarized Percentage of MAT 7 Stanine Scores for
Reading Recovery Group**

Category	1994-95		1995-96		1996-97		1997-98	
	Number of Participants	%						
below average (stanine 1-3)	38	55.9	28	42.4	29	48.3	23	40.4
average (stanine 4-6)	26	38.3	36	54.5	28	46.6	31	54.3
above average (stanine 7-9)	4	5.9	2	3.0	3	5.0	3	5.3
Total	68	100.1	66	99.9*	60	99.9*	57	100.0

*Total does not equal 100 due to rounding.

BEST COPY AVAILABLE

Although the percent of students in the “above average” category remained relatively stable from 1995 through 1998, a “see-saw” effect was exhibited at the “average” and “below average” category.

In spring 1995, 55.9% of the students were in the below average group. The percentage dipped in the spring 1996 to 42.4%. In the spring 1997, it rose to 48.3% and was followed by a decrease to 40.4% in spring 1998. Similarly, the percent of students in the “average” category had yearly fluctuations. These movements over time suggest that factors other than Reading Recovery participation may be responsible.

The Reading Recovery group did show reading improvement throughout the study. Using 1994-95 as the base year, MAT7 data from frequency distribution for spring 1996 indicate that 54.5% of the students scored in the average range, an increase of 16.2% over the spring 1995 value, while 3.0% scored in the above average range, which reflects a decrease of 2.9%. In the spring 1997, 46.6% scored in the average range, an increase of 8.3% over the spring 1995 value, while 5.0% scored in the above average range, which reflects a decrease of 0.9%. For spring 1998, 54.3% scored in the average range, an increase of 16.0% over the spring 1995 value; 5.3% scored in the above average range, reflecting a decrease of 0.6%. The Reading Recovery students did not score consistently on the MAT7 over the four-year time frame. However, they did show improvement in reading each year when compared to the spring 1995 MAT7 results.

In the spring of 1993, the **California Achievement Test, Form E (CAT-E)** was administered to kindergarten students. Out of 135 kindergarten students involved in the study, 114 students took the test. Sixty-six were in the Reading Recovery group and forty-eight in the control group. None of the 114 students had exposure to Reading Recovery during the kindergarten year. During the year 1992-1993, their total reading scale scores on the CAT-E are summarized in **Table 10**.

Data indicate that the Reading Recovery group scored slightly above the control group in total reading on the CAT-E test administered at the kindergarten level. For total reading, the Reading Recovery group had a mean scale score of 531.47, and the control group had a mean scale score of 530.69. The difference of 0.78 indicates that the Reading Recovery group scored slightly higher in total reading. This difference is *not* significant ($t = .055$).

Table 10
Summary of CAT-E Total Reading Scale Scores
Spring, 1993
(Kindergarten)

<i>Category</i>	<i>Number of Participants</i>	<i>Mean</i>	<i>Std. Deviation</i>
Reading Recovery Total Reading	66	531.47	63.61
Control Group Total Reading	48	530.69	87.38

During the spring of 1994, 122 of the 135 students involved in the study took the California Achievement Test, Form E (CAT-E) as first grade students. Out of 122 students, seventy students were in the Reading Recovery Program during the 1993-94 school year, and fifty-two were from the control group with no exposure to Reading Recovery during 1993-94. Their vocabulary, comprehension and total reading scale scores on the CAT-E are summarized in Table 11.

Data indicate that the Control Group outscored the Reading Recovery group in both Vocabulary and Total Reading on the CAT-E test administered in Grade 1. For Vocabulary, the Reading Recovery group had a mean scale score of 506.40, and the control group had a mean scale score of 523.08. The difference of 16.68 indicates that the control group scored higher in vocabulary. This difference is *not* significant ($t = -1.009$). For Comprehension, the Reading Recovery group had a mean scale score of 527.89, and the control group had a mean scale score of 522.73. The difference of 5.16 indicates that the Reading Recovery group scored higher in comprehension. This difference is *not* significant ($t = .260$). For Total Reading, the Reading Recovery group had a mean scale score of 517.36, and the control group had a mean scale score of 526.33. The difference of 8.97 indicates that the control group scored higher in Total Reading. This difference is *not* significant ($t = -.546$).

Table 11
Summary of CAT-E Reading Scale Scores
Spring, 1994
(Grade 1)

<i>Category</i>	<i>Number of Participants</i>	<i>Mean</i>	<i>Std. Deviation</i>
Reading Recovery Vocabulary	70	506.40	64.60
Control Group Vocabulary	53	523.08	116.72
Reading Recovery Comprehension	70	527.89	87.90
Control Group Comprehension	52	522.73	130.63
Reading Recovery Total Reading	70	517.36	69.33
Control Group Total Reading	52	526.33	111.63

When comparing the two groups and their relationship to the normal curve distribution in both the spring of 1993 (kindergarten) CAT-E and spring of 1994 (Grade 1) CAT-E, the Reading Recovery group had less variation in scores than the control group as indicated by the smaller value of the standard deviation. This means that the Reading Recovery group's scores were more closely centered around the mean with fewer extreme scores at either the high or low

end than the control group. This implies that the reading instruction and curriculum was more consistent for the Reading Recovery group.

For both years, 1992-1993 and 1993-1994, the results on CAT-E can be interpreted to mean that although the Reading Recovery students received more uniformity in their instruction than the control group, there was no significant difference in the reading achievement of the two groups in the kindergarten, as well as, in the first grade.

The reading scores of the Reading Recovery group and the control group on the MAT7 were compared for each year from spring 1995 through spring 1998. Scale scores for the Total Reading section of the MAT7 were obtained and analyzed. Although there were seventy students in the Reading Recovery group and sixty students in the control group, fewer and fewer students took the MAT7 with each succeeding year. t-Test statistics were computed for each year.

The spring of 1995 MAT7 Total Reading scale scores are summarized in Table 12 for the Reading Recovery and control groups. The Reading Recovery group had a mean scale score of 516.59 and the control group had a mean scale score of 531.34. The mean difference between the two groups was 14.75 in favor of the control group. This difference is *not* significant at the 95% confidence level ($t(137) = -0.842$).

Table 12
Spring, 1995
MAT7 Total Reading

Group	Number of Participants	Mean	Standard Deviation
Reading Recovery	68	516.59	97.35
Control Group	50	531.34	89.38

The spring of 1996 MAT7 Total Reading scale scores are summarized in Table 13 for the Reading Recovery and control groups. The Reading Recovery group had a mean scale score of 551.21 and the control group had a mean scale score of 561.40. The mean difference between the two groups was 10.18 in favor of the control group. This difference is *not* significant at the 95% confidence level ($t = 0.862$).

Table 13
Spring, 1996
MAT7 Total Reading

Group	Number of Participants	Mean	Standard Deviation
Reading Recovery	66	551.21	75.88
Control Group	48	561.40	35.93

The spring of 1997 MAT7 Total Reading scale scores are summarized in Table 14 for Reading Recovery and control groups. The Reading Recovery group had a mean scale score of 579.48 and the control group had a mean scale score of 588.04. The mean difference between the groups was 8.56 in favor of the control group. This difference is *not* significant at the 95% confidence level ($t = -1.244$).

Table 14
Spring, 1997
MAT7 Total Reading

Group	Number of Participants	Mean	Standard Deviation
Reading Recovery	60	579.48	33.58
Control Group	45	588.04	36.56

The spring of 1998 MAT7 Total Reading scale scores are summarized in Table 15 for Reading Recovery and control groups. The Reading Recovery group had a mean scale score of 609.26 and the control group had a mean scale score of 604.51. The mean difference between the groups was 4.75 in favor of the Reading Recovery group. This difference is *not* significant at the 95% confidence level ($t = 0.689$).

Table 15
Spring, 1998
MAT7 Total Reading

Group	Number of Participants	Mean	Standard Deviation
Reading Recovery	57	609.26	34.07
Control Group	45	604.51	35.21

An alternate view of the data was obtained by converting the MAT7 scale scores into stanines and then categorizing the stanines into three categories: (1) "below average" (*stanines 1-3*), (2) "average" (*stanines 4-6*), and (3) "above average" (*stanines 7-9*). Frequency distributions were prepared for each year of the study. A Chi-square test of significance was conducted on the distribution.

BEST COPY AVAILABLE

Table 16 outlines frequency distributions of the spring 1995 MAT7 total reading data for both groups. The control group had 15.9% fewer students at the below average range; 7.8% more students at the average range; and 8.1% more students at the above average range when compared with the Reading Recovery group. This represents a significant difference between the two groups, $X^2(2, N = 118) = 17.19, p < .05$. The control group had 60% of its students at the average to above average level of achievement compared to 44.1% of the Reading Recovery group.

Table 16
Number and Percent of Students in Various Categories of Achievement
MAT7 Total Reading
Spring, 1995

Category of Achievement	Reading Recovery		Control		Difference
	<i>Number</i>	<i>%</i>	<i>Number</i>	<i>%</i>	<i>%</i>
Below Average	<i>38</i>	<i>55.9</i>	<i>20</i>	<i>40.0</i>	<i>-15.9</i>
Average	<i>26</i>	<i>38.2</i>	<i>23</i>	<i>46.0</i>	<i>7.8</i>
Above Average	<i>4</i>	<i>5.9</i>	<i>7</i>	<i>14.0</i>	<i>8.1</i>
Total	<i>68</i>	<i>100.0</i>	<i>50</i>	<i>100.0</i>	

}

Table 17 outlines frequency distributions of the spring 1996 MAT7 total reading data of both groups. The control group had 0.7% fewer students at the below average range; 4.5% fewer students at the average range; and 5.3% more students at the above average range when compared against the Reading Recovery group. This represents a significant difference between the two groups, $X^2(2, N = 112) = 9.75, p < .05$. The control group had 58.3% of its students at the average to above average level of achievement compared to 57.5% of the Reading Recovery group.

Table 17
Number and Percent of Students in Various Categories of Achievement
MAT7 Total Reading
Spring, 1996

Category of Achievement	Reading Recovery		Control		Difference
	Number	%	Number	%	%
Below Average	28	42.4	20	41.7	-0.7
Average	36	54.5	24	50.0	-4.5
Above Average	2	3.0	4	8.3	5.3
Total	66	100.0	48	100.0	

BEST COPY AVAILABLE

Table 18 outlines frequency distributions of the spring 1997 MAT7 total reading data of both groups. The control group had 8.3% fewer students at the below average range; 8.9% fewer students at the average range; and 17.2% more students at the above average range when compared against the Reading Recovery group. This represents a significant difference between the two groups, $X^2(2, N = 105) = 62.29, p < .05$. The control group had 60% of its students at the average to above average level of achievement compared to 51.7% of the Reading Recovery group.

Table 18
Number and Percent of Students in Various Categories of Achievement
MAT7 Total Reading
Spring, 1997

Category of Achievement	Reading Recovery		Control		Difference
	Number	%	Number	%	%
Below Average	29	48.3	18	40.0	-8.3
Average	28	46.7	17	37.8	-8.9
Above Average	3	5.0	10	22.2	17.2
Total	60	100.0	45	100.0	

BEST COPY AVAILABLE

Table 19 outlines frequency distributions of the spring 1998 MAT7 total reading data of both groups. The control group had 2.6% fewer students at the below average range; 5.6% more students at the average range; and 3.1% fewer students at the above average range when compared against the Reading Recovery group. These results are *not* significant, $X^2 (2, N = 102) = 2.56, p > .05$. The control group had 62.2% of its students at the average to above average level of achievement compared to 59.7% of the Reading Recovery group.

Table 19
Number and Percent of Students in Various Categories of Achievement
MAT7 Total Reading
Spring, 1998

Category of Achievement	Reading Recovery		Control		Difference
	Number	%	Number	%	%
Below Average	23	40.4	17	37.8	-2.6
Average	31	54.4	27	60.0	5.6
Above Average	3	5.3	1	2.2	-3.1
Total	57	100.0	45	100.0	

Fifty-five Reading Recovery students and forty-one control group students took the spring of 1997 Grade 4 MEAP. The percent of students in each category of achievement of the MEAP were compared. A summary of the results is displayed in Table 20.

Only about one in five of the control group (19.5%) performed at the "Low" category of achievement on the MEAP. This compares with about two of five Reading Recovery students (41.8%) scoring in the "Low" category. Over half (53.7%) of the control group students scored in the "Satisfactory" category, compared to 38.2% of the Reading Recovery students. There is a significant difference between the Reading Recovery group and the control group in the percent of students in each category of achievement on the MEAP in favor of the control group, $X^2 (2, N = 96) = 20.50, p < .05$.

Table 20
1996-97 Grade 4 MEAP Reading

Category of Achievement	Reading Recovery		Control		Difference
	Count	%	Count	%	%
Low	23	41.8	8	19.5	-22.3
Moderate	11	20.0	11	26.8	6.8
Satisfactory	21	38.2	22	53.7	15.5
Total	55	100.0	41	100.0	

Conclusion

Is Reading Recovery successful? The Reading Recovery group, when shown separately from the control group, did show reading improvement throughout this study. Using 1994-95 as the base year, MAT7 data from frequency distribution for 1995-96 indicate that 54.5% of the students scored in the average range, an increase of 16.2% over the 1994-95 value, while 3.0% scored in the above average range, which reflects a decrease of 2.9%. In 1996-97, 46.6% scored in the above average range, an increase of 8.3% over the 1994-95 value, while 5.0% scored in the above average range, which reflects a decrease of 0.9%. For 1997-98, 54.3% scored in the average range, an increase of 16.0% over the 1994-95 value; 5.3% scored in the above average range, reflecting a decrease of 0.6%. The Reading Recovery students did not score consistently on the MAT7 over the four-year time frame; however, they did show improvement in reading each year when compared to the 1994-95 MAT7 results.

On the other hand, in many instances there were significant differences in the statistical findings between the Reading Recovery and control group when compared. Both groups performed around the same level of achievement on the California Achievement Test – Form E (CAT-E) for both the spring 1993 (kindergarten level), as well as, the spring 1994 (Grade 1). The difference in performance occurred on the Metropolitan Achievement Test (MAT7) from spring of 1994 to spring of 1998, and on the spring of 1997 MEAP.

The lack of a difference in the scores of the two groups on the spring 1993 CAT-E indicate that the control group used for the study was a valid comparison group. Unfortunately, this also indicates that the screening of students most in need of supplemental instruction may need some revision. The students in the control group were identified by classroom teachers as performing slightly above the children selected for Reading Recovery. Based upon the kindergarten CAT-E scores, this was not necessarily the case. In fact, the mean score of the Reading Recovery group was higher than the control group's mean score. However, the Reading Recovery group does appear to be more homogeneous as reflected by the smaller value of their standard deviation of scores. This could be explained by the use of Clay's Six Assessment Tasks observation survey to further screen those students identified through classroom teacher recommendation.

Similarly, the comparison of the spring 1994 CAT-E scores, when the students were in first grade, shows no significant difference. This is contrary to what would be expected since the students in the experimental group had received up to one year of Reading Recovery instruction. However, the Reading Recovery group had less variation in scores than the control group as indicated by the smaller value of the standard deviation. This implies that the reading instruction and curriculum was more consistent for the Reading Recovery group.

The mean scale score results on the MAT7 showed that the control group outperformed the Reading Recovery group in three of the four years of the study. Only in 1997-98 did the Reading Recovery group have a higher mean scale score.

Upon examining the MAT7 stanine scores and related categories of achievement, a significant difference existed between the two groups for each year 1994-95 through 1996-97, in favor of the control group. For 1997-98, the percent of students in the average and above average categories was higher for the control group although the difference was not significant.

Examination of the 1996-97 MEAP revealed similar results; the control group outperformed the Reading Recovery group. The difference between the two groups was significant.

Based on the findings of this longitudinal study over a five-year time frame (1993-94 to 1997-98), Reading Recovery was successful in having its participants maintain their level of achievement. However, the Reading Recovery group consistently scored below the achievement levels of the control group. Often, these differences were significant.

Recommendations

Based upon this longitudinal study, the following recommendations are listed:

1. If the Reading Recovery program is continued, it must be closely monitored for documentation of academic achievement and effectiveness.
2. The initial screening process of students selected for the Reading Recovery program needs revision.

APPENDIX 1a

Reading Recovery CLAY'S SIX ASSESSMENT TASKS*

- * Letter Identification
 - * Word Test
- * Concepts About Print
- * Writing Vocabulary
- * Dictation Task – Hearing and Recording Sounds in Words
 - * Text Reading Level

* Administered by the Reading Recovery teachers to the lowest 20 – 30% of the students as ranked by Kindergarten and/or First Grade teachers.

APPENDIX 1b

CLAY'S SIX ASSESSMENT TASKS **An Observation Survey of Early Literacy Achievement***

Letter Identification

“Children are asked to identify 54 characters, the upper and lower case standard letters as well as the print form of *a* and *g*. Identification may be by letter name, sound, or a word that starts with that letter.”

Word Test

“Children read a list of frequently occurring words. Three alternative lists of 20 words are available for testing and re-testing. The test was constructed using high frequency words from the Dolch Word List.”

Concepts About Print

“The examiner reads a short book and invites children to perform a variety of tasks to find out what the child has learned about the way spoken language is put into print. Two versions are available, *Sand and Stones*. The test reflects important concepts to be acquired by children in the beginning stages of learning to read. As children move from non-reading to reading, changes occur in the scores of this measure. Some of the concepts tested are: the front of the book, that print (not the picture) tells the story, that there are letters, and clusters of letters called words, that there are first letters, and clusters of letters called words, that there are first letters and last letters in words, that you can match upper or lower case letters, that spaces are there for a reason, and that different punctuation marks have meaning. Twenty four separate concepts are evaluated.”

Writing Vocabulary

“Children are asked to write all of the words they can within a maximum 10-minute limit. Within guidelines for testing, the examiner is permitted to prompt as needed.”

Hearing and Recording Sounds in Words

“The examiner reads a short sentence or two and asks the child to write the words. Children’s scores represent every sound recorded accurately in this assessment of phonemic awareness and/or orthographic awareness. Scores range from zero to thirty-seven.”

* Description supplied by the Reading Recovery staff in the Office of Communication Arts.

APPENDIX 1c

CLAY'S SIX ASSESSMENT TASKS **An Observation Survey of Early Literacy Achievement***

Text Reading

“Children are asked to read a series of increasingly more difficult texts that they have not seen before. The tester provides a minimal, scripted introduction and records reading behaviors using a running record. The texts used for Reading Recovery testing in the U.S. are not used in instruction nor were they created for Reading Recovery. Texts were drawn from established basal systems and have, over the years, been shown to be a stable measure of reading performance. Texts represent escalating gradients of difficulty.”

Selection of Students

“The children’s raw scores are recorded for each tasks. The children receiving the lowest overall scores are selected for Reading Recovery.”

* Description supplied by the Reading Recovery staff in the Office of Communication Arts.



U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)



NOTICE

REPRODUCTION BASIS



This document is covered by a signed “Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a “Specific Document” Release form.



This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either “Specific Document” or “Blanket”).