DOCUMENT RESUME

ED 293 085 CS 009 008

AUTHOR Halfar, Christine; Collins, Carla
TITLE Sustained Effects of Chapter 1 Participation: 1983-1985. OEA Analytic Report.
INSTITUTION New York City Board of Education, Brooklyn. Office of Educational Assessment.
PUB DATE Oct 87
NOTE 49p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Elementary Secondary Education; *Program Evaluation; *Reading Achievement; *Reading Improvement; *Reading Instruction; *Reading Programs; Reading Research; *Reading Tests
IDENTIFIERS Chapter 1 Programs; *Sustaining Effects Study

ABSTRACT

Conforming to the U.S. Department of Education's Chapter 1 Evaluation and Reporting System which requires local education agencies to report periodically on the sustained effects of participation in Chapter 1 programs, a sustained effects study evaluated student performance in New York City's Chapter 1 English-language reading programs during the 1983-84 school year. Student achievement was examined at three points: spring 1983 (pretest), spring 1984 (posttest), and spring 1985 (sustained effects test), using the California Achievement Test (CAT) at all three points. Data were available for 15,541 students who were in grades 3 through 7 in 1983-84. Findings indicated that average reading scores increased during the year of Chapter 1 service and also during the year following the end of program participation. Initial gains were greatest for younger students. For all but grade 3, which showed the largest increase, gains during the follow-up year were uniform. A comparison of the achievement of students who continued in Chapter 1 during the 1984-85 school year with those who did not showed that students who remained eligible for Chapter 1 services had slightly lower pre-test scores and showed little or no gains during the year of program participation. However, during the first year of Chapter 1 service they showed considerable gains. (Four appendixes include information on the creation of data files, data loss, and statistical adjustments, and a description CAT including a definition of test scores used. (SKC)

Analytic Report
O.E.A. Data Analysis Section
Raymond Domanico, Administrator
October, 1987

SUSTAINED EFFECTS OF CHAPTER 1 PARTICIPATION 1983 - 1985

Prepared by the

Analytic Studies Unit

Christine Halfar,
Manager

Carla Collins,
Analyst

New York City Board of Education
Office of Educational Assessment
Richard Guttenberg, Director

It is the policy of the Board of Education not to discriminate on the basis of race, creed, national origin, age, handicapping condition, sexual orientation, or sex, in its educational programs, activities, and employment policies, as required by law. Any person who believes he or she has been discriminated against should contact: Carole Guerra, Local Equal Opportunity Coordinator, Office of Educational Assessment, 110 Livingston Street, Room 743, Brooklyn, New York 11201. Inquiries regarding compliance with appropriate laws may also be directed to: Mercedes A. Nesfield, Director, Office of Equal Opportunity, 110 Livingston Street, Room 601, Brooklyn, New York; or the Director, Office of Civil Rights, U.S. Department of Education, 26 Federal Plaza, Room 33-130, New York, New York 10278.

Summary of the Report

According to State Education Department records, during the 1983-84 school year over 70 thousand New York City students in grades one through nine participated in English-language reading programs funded by Chapter 1 and administered, for the most part, by the 32 community school districts. The actual instructional programs varied from district to district on such dimensions as subject area, amount of instruction, organization of service delivery, and group size. (An additional 70 thousand community school district students, including kindergarten students, received Chapter 1 services in subject areas such as readiness, mathematics, English as a second language, and reading in a native language).

This report addresses the major question: across all the different Chapter 1 English-language reading programs, were the effects of program participation sustained in the year following service?

Following federal guidelines for conducting such a study, student achievement at three points was examined: spring, 1983, the pre-test; spring, 1984, the post-test; and spring, 1985, the sustained effects test. Reading achievement was measured through the citywide achievement test which, for these three years, was the California Achievement Test (CAT). The study was limited to students in grades three through seven who had valid test scores and who were promoted to the next grade each year. Data were available for 15,541 students; 10,181 received Chapter 1 service in 1983-84 and not in 1984-85, and 5,360 received Chapter 1 service both years.

Findings indicated that the effects of Chapter 1 participation were sustained during the year following service. Average reading scores increased by 1.9 normal curve equivalent units (N.C.E.) during 1983-84, the program year, and increased an additional 2.0 N.C.E. units during 1984-85, following Chapter 1 participation. Gains during the program year were greatest for younger students, a finding which parallels national findings. Gains during the follow-up year were fairly uniform across grades.

Data for the students who continued to receive Chapter 1 service in 1984-85 showed that they had slightly lower pre-test scores and showed a slight decline of 0.2 N.C.E. units in reading scores during the first year of program participation. However, they showed a substantial gain of 3.4 N.C.E. units during the second year of service. Overall gains from 1983 to 1985 were comparable for the two groups.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

I. INTRODUCTION

The United States Department of Education's Chapter 1 Evaluation and Repo·ting System requires local education agencies to report periodically on the sustained effects of participation in Chapter 1 programs. In order to meet this requirement, individual student performance must be evaluated at three points in time: before participation in Chapter 1, or the pre-test; after participation, the post-test; and one year later, the sustained effects test. In New York City each community school district is responsible for evaluating its own Chapter 1 program. However, the districts have had great difficulty conducting sustained effects studies because of high student mobility citywide; they are unable to track students who move out of their districts. The Office of Educational Assessment (O.E.A.) has access to centrally-maintained data files which permit the tracking of students as long as they remain within the city school system.

This report presents the findings of a study requested by the Office of Funded Programs and conducted by O.E.A. which uses these data files to analyze the sustained effects of 1983-84 Chapter 1 program participation. The April, 1983 citywide administration of the California Achievement Test (CAT) provided students' pre-test scores; the April, 1984 test provided post-test scores; and the April, 1985 test was the sustained effects test.

The study was limited to participants who were in grades three through seven in 1983-84. Because of the structure of the citywide testing program, these were the only students who could have been tested all three years.

The report addresses the following descriptive questions:

- How many students participated in Chapter 1 English-language reading programs in 1983-84? How many of these were in grades three through seven? In which programs did they participate?
- How many participants in grades three through seven had valid CAT scores in 1983, in 1984, or in 1985? How many had valid CAT scores all three years? If there was no CAT score, to what extent was this because students were discharged? How many students were held over, and therefore received the same level test more than once?
- How many 1983-84 participants continued to receive Chapter 1 reading services in 1984-85?
- How many participants each year received multiple Chapter 1 reading services?
- How many hours of instruction did Chapter 1 participants receive? What models of service delivery were used?

The study addresses the following analytic questions on sustained effects:

- What were the reading scores of 1983-84 Chapter 1 participants in 1983, 1984, and 1985?
- Were the effects of Chapter 1 participation sustained during the year following the end of service?
- Was there variability across grades in the degree to which Chapter 1 effects were sustained?

The following supplementary question is also addressed:

- Were there differences in the amount and pattern of gains for students who continued in Chapter 1 in 1984-85 versus those whose Chapter 1 service ended in 1983-84?

Section II presents the findings, including the description of program services and the analyses of sustained effects. For clarity, the tables appear together at the end of the chapter. Section III presents the conclusions. Description of the CAT and its test scores, creation of data files, and analytic procedures used appear in the Appendices.

II. FINDINGS

PROGRAM DESCRIPTION

According to State Education Department records, a total of 70,264 community school district students in grades one through nine received Chapter 1 English-language reading services during 1983-84. Subject areas included: reading, for 76 percent of the students; reading in English for bilingual students, for 18 percent; and reading in English as a second language, for six percent. About 17 percent of the participants were reported as having limited proficiency in English (LEP). Approximately 28 percent of the 1983-84 participants continued to receive Chapter 1 service in 1984-85. An additional 70,000 community school district students, including kindergarten students, received Chapter 1 services in subject areas such as readiness, mathematics, English as a second language, and reading in a native language. (See Appendix A for further data on all participants in Chapter 1 reading programs.)

According to data reported in their funding proposals, 28 of the 32 districts selected Chapter 1 reading program participants on the basis of their scores on the citywide reading test, the CAT. Most districts also used the CAT to evaluate program effectiveness.

Much Chapter 1 instruction was provided in small group settings. In 16 districts, proposed instructional group size was 10 students or fewer, and in the other 16 districts group

size was generally 11 to 15 students.

Most districts, 23 out of 32, used either learning stations or learning laboratories for instruction. In a learning station, remedial instruction using printed and other instructional materials is provided under professional supervision outside the regular classroom. In a learning laboratory, educational technology such as computers are used. In both modes, all Chapter 1 instruction is coordinated with the instruction by the regular classroom teacher. In the other nine districts instruction was provided in the regular classroom by the regular classroom teacher together with other professional or non-professional staff.

Sessions were generally 35 to 45 minutes long, and most districts held four or five sessions per week. Seven districts had three sessions per week and two had more than five. In all but three districts, the program operated for at least 36 weeks during 1983-84. Across most districts, then, the total number of hours of program service possible ranged from about 80 to about 125 hours over the course of the school year.

SUSTAINED EFFECTS SAMPLE

The 42,587 Chapter 1 students who were in grades three through seven in 1983-84 comprised the basic sample for study, as these were the students who participated in the citywide testing program in 1983, 1984, and 1985. About

three-fourths had valid citywide test scores for each of the three years; however, only about half of these had scores for all three years. The basic sample decreased still further with the stipulation that the three test levels be in sequence; students who had not been promoted were excluded from the study so as not to confound the results with the effects of being held over. (Analyses indicated that exclusion of holdovers did not introduce systematic bias.)

In all, there were 10,181 participants in 1983-84 Chapter 1 English-language reading programs whose CAT scores met all three requirements. These comprised the sustained effects sample. An additional 5,360 had valid test scores but because they continued to receive service in 1984-85 they were not part of the sustained effects sample. These data are presented in Table 1.

Participation in Chapter 1 by Sample

According to individual student data, 73 percent of the students in the sustained effects sample were participating in a Chapter 1 reading program for the first time. For 18 percent, 1983-84 was their second year and the remaining nine percent were reported as having had two or more years of previous experience in Chapter 1.

The distribution of service delivery models reported for individual students paralleled data from the district proposals. About 47 percent of the students in the sustained effects sample received Chapter 1 instruction outside the

regular classroom in a learning station or learning laboratory. For about 40 percent Chapter 1 instruction was provided in the regular classroom, as coordinated by the classroom teacher with other professional or non-professional staff.

Total hours of instruction over the 1983-84 school year ranged from one to over 150 hours. Over half of the students received between 50 and 100 hours, and an additional 21 percent received 126 to 150 hours. (See Table 2.)

On all descriptors of program participation the three groups, i.e., full population, the sustained effects sample, and the continuing students, showed highly similar patterns of participation. These data indicated that the two sub-groups were representative with respect to these variables.

SUSTAINED EFFECTS OF CHAPTER 1 PARTICIPATION

In order to determine whether the effects of 1983-84 Chapter 1 participation were sustained, the 1983, 1984, and 1985 CAT scores of students who did not continue to receive service in 1984-85 were examined. Gains between 1983 and 1984 represent effects of participation in Chapter 1 and gains between 1984 and 1985 comprise the sustained effects of Chapter 1 participation. The data showed that students made gains over the year of Chapter 1 participation and that these gains were sustained in the year following Chapter 1 service.

Data presented in Table 3 and Figure 1 show the mean 1983, 1984, and 1985 scores for each grade. Following the

procedures used in many federal reports, the data are reported in normal curve equivalent units (N.C.E.s), a metric which allows comparison across test levels. (See Appendix C for a further description of scores on the CAT.)

Because nearly all districts used the CAT in selecting students for Chapter 1 participation, pre-test scores were adjusted to account for regression to the mean, a statistical artifact which results from using the same test for pupil selection and program evaluation. Failure to use separate test instruments for selection and evaluation can lead to an overestimation of program impact. (See Appendix D for further information and computation of the adjusted scores.)

Across all grades, students gained an average of 1.9 N.C.E. units during the 1983-84 year of program participation and an additional 2.0 N.C.E. units during the following year. It should be noted that the N.C.E. metric is such that meaningful progress is reflected from only modest increases in N.C.E. units.

There was some variability in N.C.E. gains across grades; students in grades three, four, and five gaining more during the year of Chapter 1 service than students in grades six and seven. During the follow-up year students who were in grade three when they received Chapter 1 service showed the greatest sustained effects gains.

Table 4 presents these data in scale scores. Appendix D shows the computation of the adjustment for regression to the mean.

Overall, the results showed that students who participated in Chapter 1 reading programs in 1983-84 improved their relative standing in relation to that of the norming group. This improvement was maintained and slightly enhanced during the follow-up year in which they no longer received Chapter 1 service. Accordingly, we may conclude that the effects of Chapter 1 were sustained.

GAINS OF CONTINUING CHAPTER 1 PARTICIPANTS

In order to determine whether there were differences in the amount and pattern of gains for students who continued in Chapter 1 in 1984-85, their scores were compared with those of students who did not continue. The continuing group were students who remained eligible for Chapter 1 services; accordingly, their 1984 scores were substantially lower than the non-continuing group. At the same time, their pretest scores were somewhat lower. In order to control for these initial differences between the two groups, an adjustment was made to the 1984 scores which partialled out the effects of the 1983 scores. (See Appendix D for further information and computation of the adjusted scores.)

Data presented in Table 5 and Figure 2 show the comparison between the N.C.E. scores of the students who continued and those who did not continue in Chapter 1 reading programs in 1984-85. Across all grades, continuing students showed a slight decline in N.C.E. scores between 1983 and 1984; only students in grades four and seven showed any

improvement relative to the national norming group. However, in the second year of service, the continuing group showed gains of up to nearly six N.C.E. units. Across all grades the continuing group gained 3.4 N.C.E. units between 1984 and 1985. In contrast, the students who did not continue to participate in Chapter 1 reading programs showed gains during the year of participation and also during the follow-up year.

Table 1

Number of Chapter 1 Participants
with Citywide Test Scores, by Grade
1983-85

Grade in 83-84	Number with CAT Scores for: 1983	1984	1985	Number with three CAT Scores	Number with Sequential Levels
3	5,788	6,616	6,529	3,914	2,959
4	7,064	8,088	7,829	5,081	3,761
5	5,627	5,980	5,928	4,036	3,178
6	5,443	5,810	5,494	3,704	3,073
7	5,295	5,414	5,349	3,391	2,570
Total	29,217	31,908	31,129	20,126	15,541

- Approximately 30 thousand of the 1983-84 Chapter 1 participants had CAT scores in 1983, 1984, or 1985. However, only 20 thousand had valid scores all three years and nearly 5,000 of these were held over and thus did not receive sequential levels of the test.

Table 2

Total Hours of Instruction Received by Sample

Number of Hours	Number of Students	Percentage of Population	Cumulative Percentage
25 or Less	341	3.3	3.3
26-50	1,045	10.3	13.6
51-75	2,185	21.5	35.1
76-100	3,172	31.2	66.3
101-125	939	9.2	75.5
126-150	2,084	20.5	96.0
More than 150	415	4.1	100.1
	10,181		

NOTE: Total exceeds 100 percent because of rounding.

- Over half the sample received from 50 to 100 hours of Chapter 1 instruction over the course of the 1983-84 school year.

Table 3

Pre-test, Post-test, and Sustained Effects N.C.E. Scores
of 1983-84 Chapter 1 Participants
Who Did Not Continue in 1984-85

Grade in 83-84	N	Mean N.C.E. Score			Difference		
		1983	1984	1985	83-84	84-85	83-85
3	1,592	36.4	38.4	42.5	2.0	4.1	6.1
4	2,469	39.6	45.2	45.7	5.6	0.5	6.1
5	2,113	41.3	43.6	44.1	2.3	0.5	2.8
6	2,089	42.5	43.0	44.1	0.5	1.1	1.6
7	1,918	41.9	43.6	44.1	1.7	0.5	2.2
Total	10,181	40.4	42.3	44.3	1.9	2.0	3.9

NOTE: Mean 1983 scores reflect adjustment for regression to the mean.

- During the year of program service, students in grades three, four, and five showed the greatest gains. The greatest gains during the follow-up year were shown by students who had been in grade three when they participated in Chapter 1.

Figure 1

1983, 1984, and 1985 N.C.E. Scores of Chapter 1 Participants
Who Did Not Continue to Receive Service in 1984-85

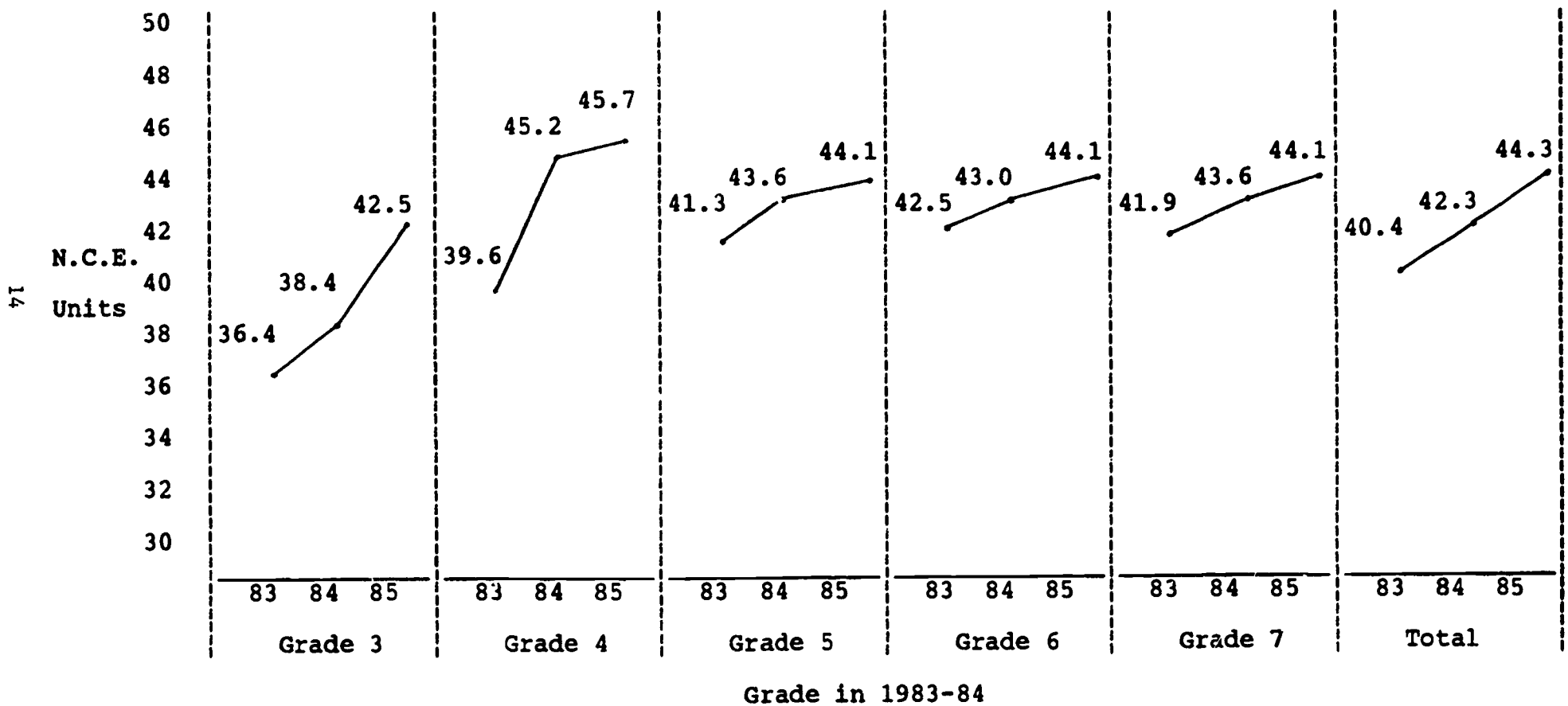


Table 4

Pre-test, Post-test, and Sustained Effects Scale Scores
of 1983-84 Chapter 1 Participants
Who Did Not Continue in 1984-85

N=10,181

Grade in 1983-84	N	1983 Adjusted Mean	1984 Observed Mean	1985 Observed Mean
3	1,592	325.7	369.7	419.2
4	2,469	372.6	426.6	457.5
5	2,113	417.1	453.9	477.8
6	2,089	447.8	475.8	500.0
7	1,918	469.5	499.4	532.9

NOTE: Mean 1983 scores are adjusted for regression to the mean.

Table 5

Comparison of 1983, 1984, and 1985 N.C.E.Scores of
1983-84 Chapter 1 Participants
Who Continued and Did Not Continue in 1984-85

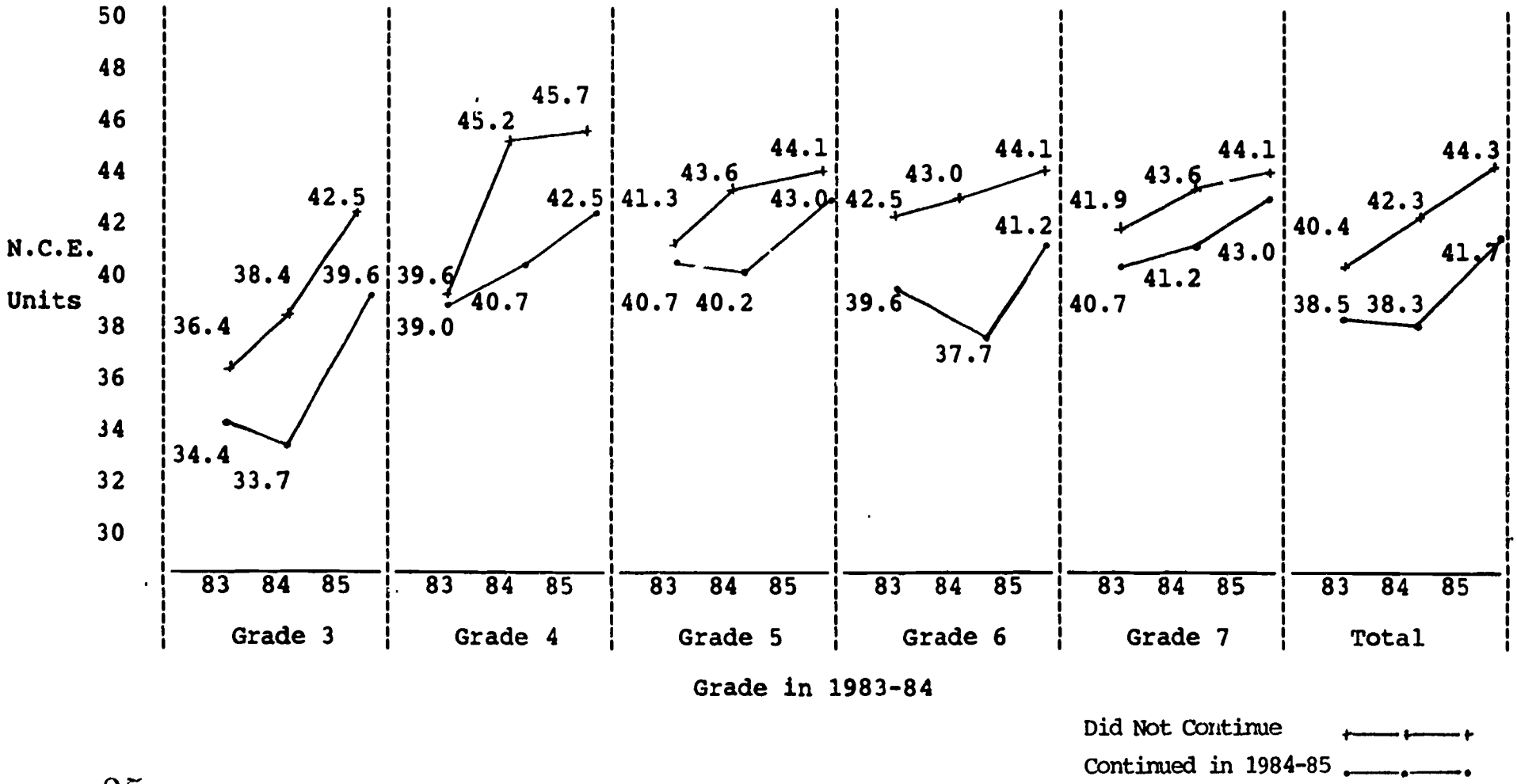
Grade	N	Continued	Mean N.C.E. Score 1983	1984	1985	Difference 83-84	84-85	83-85
3	1,592	No	36.4	38.4	42.5	2.0	4.1	6.1
	1,367	Yes	34.4	33.7	39.6	-0.7	5.9	5.2
4	2,469	No	39.6	45.2	45.7	5.6	0.5	6.1
	1,292	Yes	39.0	40.7	42.5	0.7	1.8	2.5
5	2,113	No	41.3	43.6	44.1	2.3	0.5	2.8
	1,065	Yes	40.7	40.2	43.0	-0.5	2.8	2.3
6	2,089	No	42.5	43.0	44.1	0.5	1.1	1.6
	984	Yes	39.6	37.7	41.2	-1.9	3.5	1.6
7	1,918	No	41.9	43.6	44.1	1.7	0.5	2.2
	652	Yes	40.7	41.2	43.0	0.5	1.8	2.3
Total	10,181	No	40.4	42.3	44.3	1.9	2.0	3.9
	5,360	Yes	38.5	38.3	41.7	-0.2	3.4	3.2

NOTE: Mean 1983 scores reflect adjustment for regression to the mean. Mean 1984 scores reflect adjustment for the effects of 1983 scores.

- Students who continued in Chapter 1 during 1984-85 had lower pre-test scores than those who did not continue and showed little or no gains from 1983 to 1984. However, they showed substantial gains from 1984 to 1985. Overall gains from 1983 to 1985 for the two groups were comparable.

Figure 2

1983, 1984, and 1985 N.C.E. Scores of Chapter 1 Participants
Who Continued and Did Not Continue to Receive Service in 1984-85



III. CONCLUSIONS

During the 1983-84 school year over 70 thousand students participated in English-language reading programs funded by Chapter 1 and administered by the community school districts or, in some cases, centrally. Services provided varied across districts on such dimensions as subject area, service delivery model, group size, session length, number of sessions per week, and total weeks of service.

The present study addressed this major question: across all the different programs funded by Chapter 1, were the effects of Chapter 1 participation sustained in the year following service?

Findings indicated that among students for whom data were available, average reading scores increased during the year of Chapter 1 service and also during the year following the end of program participation. Initial gains were greatest for younger students, a finding which parallels that of a national study of Chapter 1 (Carter, 1983.) For all but grade three which showed the largest increase, gains during the follow-up year were fairly uniform. Gains were evident in the standing of participating students relative to the national norming group as well as in terms of scale score units.

A comparison of the achievement of students who continued in Chapter 1 during the 1984-85 school year with

those who did not, showed that students who remained eligible for Chapter 1 services had slightly lower pre-test scores and showed little or no gains during the year of program participation. However, during the first year of Chapter 1 service they showed considerable gains. Overall gains from 1983 to 1985 were comparable for the continuing and non-continuing students.

Bibliography

Carter, L. A Study of Compensatory and Elementary Education: The Sustaining Effects Study: Final Report, (January) 1983, Pacific Palisades, CA: System Development Corporation.

Pedhazur, E.J. "Analysis of Covariance: The Unit of Analysis", Multiple Regression in Behavioral Research: Explanation and Prediction, 1973 New York: Holt, Rinehart & Winston, pp. 493-550.

Roberts, A.O.H. "Regression Toward the Mean and Regression Effects Bias", New Directions for Testing and Measurement, Number 8, 1980. San Francisco: Jossey-Bass, pp. 59-82

APPENDIX A:

CREATION OF DATA FILES

CREATION OF DATA FILES

Nine data files were used to create the basic database for the sustained effects portion of the study. These included: results of the 1983, 1984, and 1985 administrations of the California Achievement Test (CAT), the July 1984 and 1985 Biofiles, the 1983-84 and 1984-85 files of all Local Education Agency Program (LEAP) data provided to the State Education Department by the individual districts, and finally, the 1983-84 and 1984-85 LEAP data provided by centrally-based funded programs. The basic file included all students who participated in any Chapter 1 English-language reading program in 1983-84. These programs included regular remedial reading, bilingual reading, and English as a Second Language reading. The table which follows shows the total number of participants in reading programs in 1983-84 by grade, and the total number continuing to receive service in 1984-85.

In addition to the individual data files, background data were obtained from the individual district proposals on file in the Office of Funded Programs. These proposals provide such information as the programs' model for instruction and criteria for selection of participants.

Participation in Chapter 1 Reading Programs

1983 - 1985

Grade in 83-84	Total Number of 83-84 Participants	Reading Program			Number Receiving Multiple Services	1983-84 Participants Continuing in 1984-85
		English	Bilingual	?.S.L.		
3	9,007	6,724	1,831	452	207	2,941
4	10,393	8,428	1,482	483	300	3,478
5	7,790	6,550	856	384	112	2,783
6	7,760	6,406	1,006	348	149	2,273
7	7,637	5,841	1,339	457	43	2,030
Total	42,587	33,949	6,514	2,124	811	13,505
1	8,739	5,935	2,164	640	132	1,062
2	9,517	6,828	1,934	755	240	2,932
8	5,439	4,171	818	450	39	1,547
9	1,896	1,452	330	106	16	669
Unknown	2,446	1,443	978	25	0	2
	28,037	19,829	6,232	1,976	427	6,212
Citywide	70,624	53,778	12,746	4,100	1,238	19,717

APPENDIX B: DATA LOSS

DATA LOSS

As reported in Chapter II, about 70 percent of the target group had CAT scores for any of the three years under study, and only 35 percent had scores all three years. It appeared that this data loss occurred largely because of missing or invalid data on the citywide test files, as opposed to students' leaving the school system. Most of the 42 thousand grade-three-through-seven Chapter 1 participants reported on the LEAP files were located on either the 1984 or 1985 Biofiles maintained by the Student Information System; these records showed that only 11 percent of the students with incomplete data had been discharged from the New York City schools.

To determine the extent to which data loss might have introduced bias, pretest scores of students with complete data were compared to those of students with fewer than three test scores. The table which follows shows 1983 scores of students with incomplete data, students with three scores on sequential test levels, and students with three scores, not on sequential test levels. This last group were generally students who had been held over. In every grade there was a slight bias apparent in favor of students with complete data.

Comparison of Pre-test Scale Scores of Students with Incomplete Data, with Three Scores on Sequential Levels, and with Three Scores not on Sequential Levels

Grade	Students with Incomplete Data			Students with Three Scores in Sequence			Students with Three Scores Not in Sequence		
	N	Mean	(S.D.)	N	Mean	(S.D.)	N	Mean	(S.D.)
3	1,592	318.0	33.3	2,959	319.8	31.4	955	317.8	31.4
4	1,983	359.7	34.7	3,761	368.9	30.6	1,320	350.7	36.6
5	1,591	407.9	29.9	3,178	410.6	25.8	858	405.5	28.5
6	1,739	432.1	32.1	3,073	438.6	29.2	631	424.1	33.1
7	1,904	456.4	36.2	2,570	461.6	30.1	821	444.0	37.5
Total	9,091	393.6	60.8	15,541	397.2	57.1	4,585	380.9	57.6

APPENDIX C:

DESCRIPTION OF CAT AND DEFINITION OF TEST SCORES USED

DESCRIPTION OF CAT AND DEFINITION OF TEST SCORES USED

The California Achievement Test (CAT) is an achievement test which measures the knowledge and skills that students have acquired in specified content areas at a certain point in time. The CAT reading test covers vocabulary, comprehension, and in some grades, phonic analysis and structural analysis. It is administered annually to all eligible students in grades two through nine.

Test Form

For the years covered by the present study, the primary test form administered was as follows: in 1983, Form D; in 1984, Form C; and in 1985, Form D. However, in 1983 Form C was given as the pretest for third-grade participants; these students were in grade two in 1983 and second-graders always receive Form C.

Test Level

Each of the grades three to seven were administered a different level of the CAT for each of the three years, 1983, 1984, and 1985.

Test Scoring

Scale scores. Scale scores are units of a single equal-interval scale that are applied across all levels of the test regardless of grade. They provide a basis for deriving other normative scores that can be used to describe test performance. They are also particularly appropriate for

statistical procedures; for example, scale scores can be added, subtracted, and averaged across test levels allowing direct comparisons among classes, schools, or districts. Scale scores may also be used to measure the growth of individual students or groups of students from year to year.

Normal curve equivalent scores (N.C.E.s) N.C.E.s have many of the same characteristics as percentile ranks but have the advantage of being based on an equal-interval scale. The scale ranges from one to 99 with a midpoint of 50 and a standard deviation of approximately 21. As a result, the difference between two successive scores has the same meaning over all parts of the scale. N.C.E.s also provide a standardized measure of a student's level of achievement in relation to the performance of a norming sample.

Conversion of Scale Scores to N.C.E.s

All statistical adjustments were made in scale score units. (See Appendix D.) In order to obtain the equivalent N.C.E., each adjusted mean scale score was converted, first to the nearest raw score, then to a national percentile score, and finally to the corresponding N.C.E.

APPENDIX D:

STATISTICAL ADJUSTMENTS

STATISTICAL ADJUSTMENTS

Statistical Adjustment for Regression to the Mean

Because a large majority of schools also used the 1983 CAT scores as part of their selection criteria for program participation it was necessary to control for potential biases arising from this type of measurement procedure. In analyses of Chapter 1 services, the amount of change in scores between each testing is used to estimate program effectiveness. A failure to separate selection and pretest instruments results in an overestimation of the program's impact. This inflated gain is known as regression to the mean.

The equation used to adjust pretest scores to account for the regression effect is taken from A.O.H. Roberts, "Regression Toward the Mean and the Regression Effect Bias" in New Directions for Testing and Measurement, Number 8, 1980 (San Francisco, Jossey-Bass), pages 59-82.

The equation is as follows:

$$\bar{X}_{cs} = \bar{X}_{s} + \frac{\sigma^2_{s}}{} (1 - \rho_{xx}) (\bar{X}_{g} - \bar{X}_{s})$$

Values for each variable and the computations for each grade are shown in the following tables.

Values of Variables for Adjustment of Pre-test Scores for Regression Effects for the Sustained Effects Sample, by Grade
N = 10,181

Variable		Grade				
		3	4	5	6	7
$\bar{X}_{cs}$	= corrected pre-test scale score of sample	325.69	372.60	417.13	447.79	469.51
$\bar{X}_{s}$	= pre-test scale score of sample	322.67	370.20	412.25	441.91	462.66
$\bar{X}_{g}$	= citywide pre-test mean scale score	362.60	405.23	439.33	485.52	510.66
σ	= standard deviation of pre-test scale scores nationally	45.50	47.40	57.00	56.90	56.50
s	= standard deviation of pre-test scale scores citywide	54.86	60.01	53.66	55.88	61.70
ϱ	= coefficient of reliability	.89	.89	.84	.87	.83

Computations of the 1983 Scale Scores
Adjusted for Regression Effects
for Sustained Effects Sample
N = 10,181

$$\bar{X}_{cs} = \bar{X}_s + \frac{\sigma^2}{s^2}(1 - \rho_{xx})(\bar{X}_g - \bar{X}_s)$$

Grade 3

$$= 322.67 + \frac{(45.5)^2}{(54.86)^2}(1 - .89)(362.60 - 322.67)$$
$$= 325.69$$

Grade 4

$$= 370.20 + \frac{(47.5)^2}{(60.01)^2}(1 - .89)\ 405.23 - 370.20)$$
$$= 372.60$$

Grade 5

$$= 412.25 + \frac{(57.0)^2}{(53.66)^2}(1 - .84)(439.33 - 412.25)$$
$$= 417.13$$

Grade 6

$$= 441.91 + \frac{(56.9)^2}{(55.88)^2}(1 - .87)(485.52 - 441.91)$$
$$= 447.79$$

Grade 7

$$= 462.66 + \frac{(56.5)^2}{(61.70)^2}(1 - .83)(510.66 - 462.66)$$
$$= 469.51$$

Values of Variables for Adjustments
of Pre-test Scores for Regression Effects
for Continuing Participants
N = 5,360

Variable			Grade				
			3	4	5	6	7
$\bar{X}_{cs}$	=	corrected pre-test scale score of sample	320.00	371.32	413.06	438.71	465.90
$\bar{X}_{s}$	=	pre-test scale score of sample	316.51	366.43	407.27	431.42	458.45
$\bar{X}_{g}$	=	citywide pre-test mean scale score	362.60	405.23	439.33	485.52	510.66
σ	=	standard deviation of pre-test scale scores nationally	45.50	47.40	57.00	56.90	56.50
s	=	standard deviation of pre-test scale scores citywide	54.86	60.01	53.66	55.88	61.70
ϱ	=	coefficient of reliability	.89	.89	.84	.87	.83

Computations of the 1983 Scale Scores Adjusted for Regression Effects for Continuing Participants, by Grade

$$\bar{X}_{cs} = \bar{X}_s + \frac{\sigma^2}{s^2} (1 - \rho_{xx}) (\bar{X}_g - \bar{X}_s)$$

Grade 3

$$= 316.51 + \frac{(45.5)^2}{(54.86)^2} (1 - .89) (362.60 - 316.51)$$

$$= 320.00$$

Grade 4

$$= 366.43 + \frac{(47.4)^2}{(60.01)^2} (1 - .89) (405.23 - 366.43)$$

$$= 371.32$$

Grade 5

$$= \quad 27 + \frac{(57.0)^2}{(53.66)^2} (1 - .84) (439.33 - 407.27)$$

$$= 413.06$$

Grade 6

$$= 431.42 + \frac{(56.9)^2}{(55.88)^2} (1 - .87) (485.52 - 431.42)$$

$$= 438.71$$

Grade 7

$$= 458.45 + \frac{(56.5)^2}{(61.70)^2} (1 - .83) (510.66 - 458.45)$$

$$= 465.90$$

Statistical Adjustment for the Effects of the Covariate

Because our sample consisted of two groups, the sustained effects sample who after one year were no longer eligible for Chapter 1 services (Group A) and the students who continued to receive Chapter 1 services for a second year (Group B), it was necessary to consider artifactual group differences in pre-test scores. This source of bias was controlled by partialling out the effects of the 1983 test scores from the 1984 test scores through an analysis of covariance for each group. Due to the nature of the data, i.e., continuous test scores, the analyses of covariance were conducted using a multiple regression technique. Adjusted 1984 test scores were used for comparison of the two groups.

The equation used to adjust post-test scores to account for the effects of the pretest was suggested by E.J. Pedhazur in a personal communication. Additional information regarding this statistical procedure is provided in E.J. Pedhazur, "Multiple Regression in Behavioral Research: Explanation and Prediction", 1982, (New York, Holt, Rinehart, and Winston), pages 493-550.

The formula is as follows:

$$Y'_{\substack{84adjc \\ group\ A}} = Y_{\substack{84act \\ group\ A}} - b\left(\bar{X}_{\substack{83r \\ group\ A}} - \bar{\bar{X}}_{\substack{83r \\ group\ A+B}}\right)$$

Values for each variable and the computations are shown in the following tables.

Values of Variables for Adjustment of Post-test Scores
for the Effects of Pre-test Scores
for the Sustained Effects Sample
N = 10,181

Variable			Grade				
			3	4	5	6	7
Y'_{84adjc}	=	corrected mean scale score of Group A	368.21	425.98	453.18	474.05	498.89
Y_{84obs}	=	observed mean scale score of Group A	369.68	426.57	453.87	475.76	499.39
b	=	weighted regression coefficient	.56	.49	.51	.59	.54
$\bar{X}_{83r}$	=	mean adjusted pre-test scale score of Group A	325.69	372.60	417.13	447.79	469.51
= X_{83r}	=	mean adjusted pre-test scale score of Groups A and B	323.06	371.40	415.77	444.89	468.59

Computation of Adjustment of Post-test Scores
for Effects of Pre-test Scores
for the Sustained Effects Sample, by Grade
N = 10,181

$$Y'_{84adjc\ group\ A} = Y_{84act\ group\ A} - b\,(\bar{X}_{83r\ group\ A} - \bar{X}_{83r\ group\ A+B})$$

Grade 3

= 369.68 - .56 (325.69 - 323.06)

= 368.21

Grade 4

= 426.57 - .49 (372.6ᵕ 371.40)

= 425.98

Grade 5

= 453.87 - .51 (417.13 - 415.77)

= 453.18

Grade 6

= 475.76 - .59 (447.79 - 444.89)

= 474.05

Grade 7

= 499.39 - .54 (469.51 - 468.59)

= 498.89

Values of Variables for Adjustment of Post-test Scores
for the Effects of Pre-test Scores
for Continuing Participants, by Grade
N = 5,360

Variable			Grade 3	4	5	6	7
Y'_{84adjc}	=	corrected mean scale score of Group B	357.50	413.11	440.43	457.99	491.72
Y_{84obs}	=	observed mean scale score of Group B	355.79	413.07	439.05	454.34	490.27
b	=	weighted regression coefficient	.56	.49	.51	.59	.54
$\bar{X}_{83r}$	=	mean adjusted pre-test scale score of Group B	320.00	371.32	413.06	438.71	465.90
= X_{83r}	=	mean adjusted pre-test scale score of Groups A and B	323.06	371.40	415.77	444.89	468.59

Computation of Adjustment of Post-test Scores
for the Effects of Pre-test Scores
for Continuing Participants, by Grade
N = 5,360

$$Y'_{84adjc\ group\ B} = Y_{84act\ group\ B} - b\,(\bar{X}_{83r\ group\ B} - \bar{\bar{X}}_{83r\ group\ A + B})$$

Grade 3

= 355.79 - .56 (320.005.69 - 323.06)

= 357.5068.21

Grade 4

= 413.07 - .49 (371.32 - 371.40)

= 413.11

Grade 5

= 439.05 - .51 (413.06 - 415.77)

= 440.43

Grade 6

= 454.34 - .59 (438.71 - 444.89)

= 457.99

Grade 7

= 490.27 - .54 (465.90 - 468.59)

= 491.72