DOCUMENT RESUME

ED 137 342 TM 006 143

AUTHOR Pelavin, Sol H.; David, Jane L.

TITLE Evaluating Long-Term Achievement: An Analysis of Longitudinal Data from Compensatory Education Programs.

INSTITUTION Stanford Research Inst., Menlo Park, Calif. Educational Policy Research Center.

SPONS AGENCY Office of the Assistant Secretary for Education (DHEW), Washington, D.C.
REPORT NO EPRC-4537-15
PUB DATE Mar 77
CONTRACT 300-76-0025
NOTE 69p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Academic Achievement; *Achievement Gains; Achievement Tests; *Compensatory Education Programs; Disadvantaged Youth; Elementary Secondary Education; *Longitudinal Studies; Program Effectiveness; *Program Evaluation; Student Testing; *Time
IDENTIFIERS Elementary Secondary Education Act Title I

ABSTRACT

Several sets of data are analyzed, all of which demonstrate that estimates of achievement rates can differ dramatically when different time periods are used for an evaluation of a compensatory education program. Specifically, it is found that including the summer months in an evaluation often considerably reduces estimates of achievement and hence monthly achievement rates. In other words, achievement gains made during the school year are not sustained, even until the next fall. These findings are consistently supported by the data regardless of the standardized achievement test used, the grade level, the subject area, or the program. An evaluation of a one-year program should at least take account of fall-to-fall achievement. For multi-year programs, evaluations should be based on a time period that includes the summer following each year of the program. Including the following summer in an evaluation will allow the conclusion to reflect the extent to which student achievement is sustained and will, therefore, provide more useful information to those concerned with improving programs. (RC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED137342

EVALUATING LONG-TERM ACHIEVEMENT: AN ANALYSIS OF LONGITUDINAL DATA FROM COMPENSATORY EDUCATION PROGRAMS

Research Report
EPRC 4537-15

Prepared for:

OFFICE OF THE ASSISTANT SECRETARY
FOR EDUCATION
DEPARTMENT OF HEALTH, EDUCATION,
AND WELFARE
WASHINGTON, D.C. 20201

CONTRACT OEC-300-76-0025

EDUCATIONAL POLICY RESEARCH CENTER

POLICY RESEARCH REPORT

A Policy Research Report is an official document of the Educational Policy Research Center. It presents results of work directed toward specific research objectives. The report is a comprehensive treatment of the objectives, scope, methodology, data, analyses, and conclusions, and presents the background, practical significance, and technical information required for a complete and full understanding of the research activity. The report is designed to be directly useful to educational policy makers.

RESEARCH MEMORANDUM

A Research Memorandum is a working paper that presents the results of work in progress. The purpose of the Research Memorandum is to invite comment on research in progress. It is a comprehensive treatment of a single research area or of a facet of a research area within a larger field of study. The Memorandum presents the background, objectives, scope, summary, and conclusions, as well as method and approach, in a condensed form. Since it presents views and conclusions drawn during the progress of research activity, it may be expanded or modified in the light of further research.

RESEARCH NOTE

A Research Note is a working paper that presents the results of study related to a single phase or factor of a research problem. It also may present preliminary exploration of an educational policy issue or an interim report which may later appear as a larger study. The purpose of the Research Note is to instigate discussion and criticism. It presents the concepts, findings, and/or conclusions of the author. It may be altered, expanded, or withdrawn at any time.

SRI Project 4537 March 1977

Educational Policy Research Center

SOL H. PELAVIN
JANE L. DAVID

Research Report

EPRC 4537-15

EVALUATING LONG-TERM ACHIEVEMENT: AN ANALYSIS OF LONGITUDINAL DATA FROM COMPENSATORY EDUCATION PROGRAMS

Prepared for:

OFFICE OF THE ASSISTANT SECRETARY
FOR EDUCATION
DEPARTMENT OF HEALTH, EDUCATION,
AND WELFARE
WASHINGTON D.C. 20201

CONTRACT OEC-300-76-0025

EXECUTIVE SUMMARY

Support for compensatory education rests on the assumption that greater achievement can improve the academic future of disadvantaged students, and that a better academic future can improve their life chances. In order to change students' futures, an increase in achievement from compensatory education should persist in the years following a student's participation in compensatory education programs. Therefore, evaluations of compensatory education should include measures of long-term, or sustained, achievement. At the very least, these measures should reflect the extent to which increases in achievement persist over the summer following a school-year program.

Evaluations of compensatory education such as those of ESEA Title I have generally been based on tests administered in the fall and spring of a given school year, rather than on tests measuring sustained achievement. The findings of SRI's study of Title I evaluations* and research on achievement over the summer suggest that high gains of disadvantaged students during the school year might be followed by large losses during the summer. Thus, evaluations reflecting the extent to which gains are sustained over the summer could produce conclusions about the success of a program very different from those based on measures of fall-to-spring achievement.

To test the hypothesis that the period of time on which an evaluation is based can significantly alter the conclusions reached, we obtained data from evaluations of several compensatory education programs. We analyzed these data by calculating estimates of achievement rates over different time periods and comparing the results.

---

*T. Thomas and S. Pelavin, "Patterns in ESEA Title I Reading Achievement," EPRC 4537-12, Stanford Research Institute, Menlo Park, California (March 1976).

Our analyses demonstrate that estimates of achievement rates differ dramatically when different time periods are used. Specifically, our results show that including the summer months in an evaluation can considerably reduce estimates of achievement. Large increases in school-year achievement are not sustained even until the next fall. The findings are remarkably consistent across achievement tests, grade levels, subject areas, and programs.

We conclude that if compensatory education programs are going to be evaluated on the basis of student achievement, it is essential to have measures of sustained achievement. Thus, an evaluation of a one-year program should at least take account of fall-to-fall achievement. For multi-year programs, we suggest that evaluations be based on a time period that includes the summer following each year of the program. Including the following summer in an evaluation will allow the conclusions to reflect the extent to which student achievement is sustained and will, therefore, provide more useful information to those concerned with improving programs.

School districts should administer achievement tests at least every fall and preferably every fall and spring. These data would show the extent to which school-year gains are sustained through the following summer. Also, the extent and causes of summer losses should be explored. Since there are no simple solutions (for example, little research exists to show that summer school would alleviate the summer losses), it is important to be able to determine why the losses occur in order to develop appropriate remedies.

CONTENTS

ILLUSTRATIONS

TABLES

TABLES (continued)

TABLES (concluded)

ACKNOWLEDGEMENTS

We were fortunate in receiving support and cooperation from a large number of people in the process of carrying out this work and preparing the report. We are deeply indebted to Susan Bryson who miraculously transformed reams of handwritten sheets of paper into computer-readable tapes and finally the tables that appear in the report. She also contributed immensely to the preparation of the report.

We would also like to acknowledge those who reviewed the draft report, including Drs. Gerry Hendrickson and Keith Baker, Office of the Assistant Secretary for Planning and Evaluation, DHEW; Elizabeth Demarest, Office of the Assistant Secretary for Education, DHEW; Dr. George Mayeske, OPBE, United States Office of Education, DHEW; and Dr. Marshall Smith, National Institute of Education, DHEW. Their reactions greatly improved the final version of the report.

Our gratitude also goes to Arlowayne Brenno, Bonnie Miller, and Jane O'Connor who suffered through the typing and copying of numerous drafts and revisions with inordinate skill, patience, and cheer.

Finally, we are indebted to Elizabeth Reisner who has served as our project monitor in the Office of the Assistant Secretary for Education. Her support, patience, and valuable criticism facilitated the progress of the study and greatly enhanced the quality of the final report. Though we acknowledge all those mentioned, the responsibility for the interpretations, conclusions, and faults are ours alone.

I INTRODUCTION

A fundamental assumption underlying compensatory education is that greater achievement can improve the academic future of disadvantaged students, and that this in turn can improve their "life chances." Therefore, one of the goals of Title I of the Elementary and Secondary Education Act (ESEA) is to increase the achievement of disadvantaged students. In order to change students' futures, this increase in achievement should be reflected in the years following students' participation in compensatory education programs. At a minimum, an increase in achievement should persist over the summer following a school-year program. However, evaluations of compensatory education in general and of ESEA Title I in particular have not included estimates of sustained achievement. Instead, they have been based on tests administered in the fall (pretest) and spring (posttest), and have used student achievement during this period, that is, the posttest score adjusted in some way for the pretest score, to judge the success of the program.

We hypothesized that evaluations based on measures of sustained achievement would lead to different conclusions than evaluations based on fall-to-spring achievement. To test this hypothesis, we obtained data from evaluations of several compensatory education programs. With these data, we demonstrate that fall and spring measures, by themselves, are insufficient indicators of program success and may even be misleading. Our conclusion is that evaluations of achievement in compensatory programs should be based on a time frame that allows estimates to be made of sustained achievement.

In addition to an appropriate time frame, one must have a procedure that can distinguish the portion of the achievement gain that is attributable to the program. There are a variety of techniques, each with its flaws, for estimating the program's effect on achievement. Arguments

over the best method, particularly in evaluations that rely on standardized achievement tests, will continue for years to come. We do not intend to argue the relative merits of these methods in this paper. Rather, our purpose is to demonstrate that regardless of how achievement is measured, the time frame used in the evaluation can dramatically affect the results.

II BACKGROUND

This work was motivated by a study of ESEA Title I achievement done by the Educational Policy Research Center (EPRC) and relates directly to that study. In "Patterns in ESEA Title I Reading Achievement," Thomas and Pelavin (1976) analyzed data from 283 State Educational Agency (SEA) Title I evaluation reports for the academic years 1968-69 through 1973-74. For those states reporting adequate data, average monthly reading achievement was calculated. These analyses showed that average monthly reading achievement across states was close to 1.1 grade-equivalent months. That is, Title I participants appeared, on the average, to be gaining more than one grade-equivalent month for each month of the school year. In terms of the unofficial standard of success, which is a month's gain for each month in the program, ESEA Title I appeared to be a success.

To corroborate this finding, the report looked for evidence of a Title I impact in the results of three statewide testing programs. Each of the three states had reported that Title I students were achieving at a monthly rate greater than 1.0. The expectation was that there would be an upward shift in the statewide test results over the years for low-percentile students (those most likely to be in Title I). The data from the statewide tests did not reflect such a shift, and hence did not provide evidence to support the apparent success of Title I as found from the SEA data.

The EPRC report offered a number of possible explanations for these apparently contradictory findings. Since the most plausible explanations involved the amount of achievement gained or lost during the summer months, a review of the relevant research was performed. The review suggested that disadvantaged students generally achieve at a much slower rate

during the summer than during the school year.* In fact, during the summer months, many disadvantaged students show no achievement gain and even suffer a loss. If this is frequently true for ESEA Title I participants, measuring achievement over a full calendar year would produce a smaller monthly rate than measuring achievement from a fall pretest to a spring posttest as in the SEA reports.

In view of the long-range goals of compensatory education, if gains made over the school year are not sustained even over the summer, program evaluations should reflect this phenomenon. The EPRC report pointed out that such speculation could be confirmed only by analyzing longitudinal data on individual students. We have obtained such data and will demonstrate in this paper how conclusions about program effectiveness change when sustained achievement is measured.

*The review includes studies based on disadvantaged students both with and without school-year compensatory education programs. The contrast is with the pattern of growth of the 50th percentile student, who is defined to achieve at a rate of nine grade-equivalent months over the school year and one month over the summer.

III METHOD OF ANALYSIS

SEA Title I evaluations almost exclusively report achievement rates based on data from a fall pretest and a spring posttest. Therefore, the evaluation reports reflect achievement during a relatively short period of time, usually seven months.* For the reasons just presented, we are particularly interested in whether or not a program's effect persists over the summer. To determine the extent to which the effect on achievement is sustained until the beginning of the next school year, data from an additional test administered in the fall of the subsequent school year were used to calculate both the fall-to-spring and the fall-to-fall rates of achievement. Comparing each of these rates to the unofficial Title I standard of success--a month-for-month achievement rate-- we established the extent to which conclusions change about the success of a program when the longer period of time is used.

The following example illustrates our analytical approach for one- and two-year evaluations. The data consist of a matched, longitudinal sample of 87 students who have participated in a compensatory education program during Grades 3 and 4.†

To evaluate one year of the program, we use grade-equivalent means for three administrations of the Reading Comprehension portion of the Gates-MacGinitie Reading Test. The means for the beginning and end of Grade 3 (3 Fall and 3 Spring) and the beginning of Grade 4 (4 Fall) are:

3 Fall	3 Spring	4 Fall
2.57	3.62	3.00

*The pretest to posttest interval does vary from four to eight months. However, the average and by far the most common is seven.

†See Table 3, line 1, in the next section.

Mean achievement from fall to spring (3 Spring minus 3 Fall) is approximately 10 months. Dividing this gain by the number of months between test administrations (approximately seven months), we get a monthly achievement rate of about 1.5 months per month. Compared to the month-for-month standard, the program would be judged successful on this basis. However, the mean achievement gain from fall to fall (4 Fall minus 3 Fall) is only four months, which translates into a monthly rate of 0.4 month per month.* During this longer time interval the program would not be judged a success.

In this example, we calculated rates of achievement for both 7-month and 12-month periods and used the month-for-month standard as the reference point for success. Month-for-month achievement is the expected rate, by definition, for the average or 50th percentile student during the school year. Though there is not an established monthly rate of achievement for the disadvantaged student, the annual rate of seven months per year is frequently used† and is supported by national Title I data.± Therefore we prefer to use an expected annual achievement rate of seven months gain as our point of comparison rather than a monthly rate of achievement. In our example, the annual achievement of four months clearly does not exceed the expected seven months and suggests that high achievement during the pretest-to-posttest period is not even sustained until the beginning of the next school year.

---

*A grade-equivalent year, by definition, consists of 10 grade-equivalent months. Therefore, to calculate a monthly rate in grade equivalents, the annual gain is divided by 10.

†The Office of Education uses this estimate. See, for example, the 1975 U.S. General Accounting Office's "Report to the Congress by the Controller General of the United States, Assessment of Reading Activities Funded Under the Federal Program of Aid for Educationally Deprived Children."

±T. Thomas and S. Pelavin, op. cit., have calculated annual achievement by dividing pretest (fall) scores for large samples of disadvantaged students by the grade level (the number of years the students had been in school). For all grades, these annual rates were close to seven months. These samples included some students who had previously participated in Title I, hence the seven months is conceivably an overestimate of annual achievement.

To illustrate a two-year evaluation, the example is extended through a second year of the program. To the three grade-equivalent means used above, we added the results of two more administrations of the Reading Comprehension subtest. The means for the beginning and end of Grade 3 (3 Fall and 3 Spring), the beginning and end of Grade 4 (4 Fall and 4 Spring), and the beginning of Grade 5 (5 Fall) are:

3 Fall	3 Spring	4 Fall	4 Spring	5 Fall
2.57	3.62	3.00	4.12	3.53

We use these five means to demonstrate how estimates of two-year achievement rates differ under three time periods. The shaded areas in Figure 1 illustrate the three different time periods.

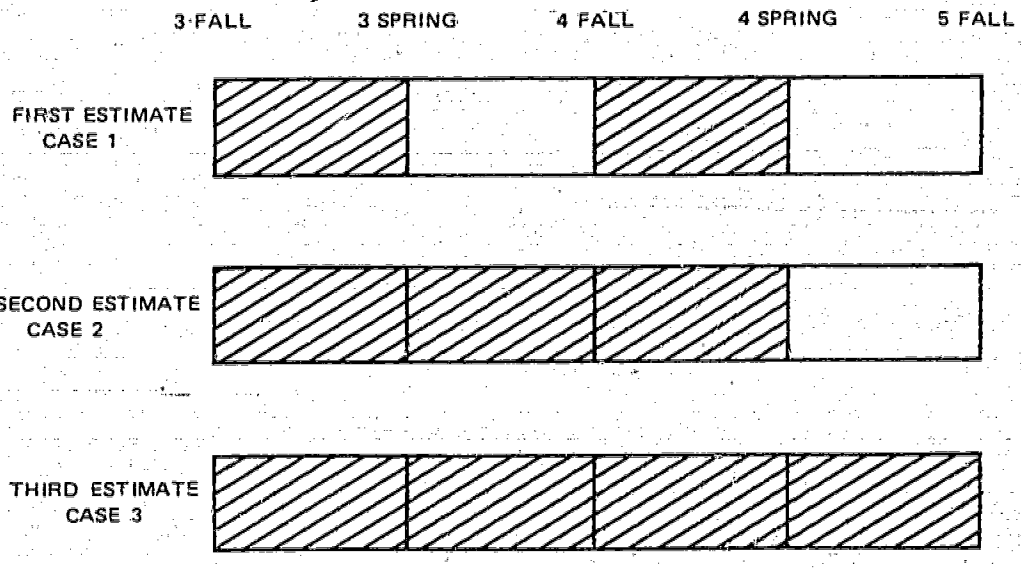


FIGURE 1 THREE POSSIBLE TIME PERIODS FOR A TWO-YEAR EVALUATION

First, we consider a two-year evaluation that estimates a monthly achievement rate for each of the pretest (fall) to posttest (spring) intervals. This is the procedure used in most multi-year Title I evaluations. We have already shown that the achievement rate from 3 Fall to

3 Spring is 1.5 months per month. Achievement for the second year (Grade 4) is approximately 12 months (4 Spring minus 4 Fall), which yields an achievement rate of 1.6 months per month. If the program were judged on the basis of these two rates, each of which exceeds the month-for-month standard, it would be considered a success.

A second estimate of the achievement rate over two years is obtained by calculating the achievement from 3 Fall to 4 Spring, thus including the intervening summer somths. The achievement during this period (4 Spring minus 3 Fall) is 15.5 months. To obtain a monthly rate, the achievement is divided by the number of months in the interval: 10 months from 3 Fall to 4 Fall + 7 months from 4 Fall to 4 Spring = 17 months. This procedure yields an achievement rate of 0.9 month per month. While this rate is considerably lower than the pretest-to-posttest rate for either Grade 3 or 4, it is nonetheless close to the month-for-month standard. Hence the program might still be judged successful.

The third time period spans two full calendar years. We subtract the Grade 3 fall test from the Grade 5 fall test (5 Fall minus 3 Fall) and divide this gain of 9.6 months by 20 months (10 months for each of the two years). This gives an achievement rate of 0.5 month per month. Clearly, for this time period the program would not be considered a success. A monthly gain of 0.5 is far below the month-for-month standard, and average achievement of five months per year is below even that expected for disadvantaged students.

Our conclusion is that the third time period is the appropriate one since it reflects the extent to which program gains persist until the beginning of the following school year. Since the gains are considerably smaller under this time period, we conclude that high gains achieved during the school year are not sustained over the summer months.

The following section describes our samples and presents the data to support these conclusions. The data presented are all in the metric of grade equivalents. We are fully cognizant of the weaknesses of the

grade-equivalent scale, particularly for measuring growth.* Nevertheless, we wanted to be consistent with previous discussions of the impact of compensatory education programs on achievement. These discussions have relied heavily on the grade-equivalent metric because it provides a means of aggregating and comparing different tests. As a result, both the standard for success and expectations for the disadvantaged student are expressed in terms of grade equivalents. However, we have carried out analyses demonstrating that the patterns of interest are not seriously affected by the metric. Since the transformations are monotonic and close to linear except at the extremes, this finding is not surprising. Appendix A presents analyses comparing grade equivalents, standard scores and percentiles.

*See, for example, Angoff (1971), Cronbach (1970), Harcourt Brace Jovanovich (1973), or Tallmadge (1975).

IV RESULTS

To compare judgments of program success under different time periods, we obtained multi-year data from two major sources: the evaluation of a large midwestern city's compensatory education program and the evaluations of four junior high school compensatory education programs in California. All the programs span several grades and collect data annually in both fall and spring. All the data had been collected originally as part of annual evaluations and were therefore cross-sectional by year.

By matching students across time and test administrations, we were able to create longitudinal data that permitted us to compare the same group of students during different time periods. The use of only matched, longitudinal samples implies that whenever two means are compared, they are based on data from exactly the same students. We present these data below. We describe first the sample and results from City M, the midwestern city, and then the samples and results from California.

City M

City M has a citywide Title I reading program for eligible students in Grades 3 through 9. The Reading Comprehension section of either Level C or D* of the Gates-MacGinitie Reading Test was administered to the students in both fall and spring of each school year. We received data for six cohorts† of students in the program during the school years 1971-72 through 1974-75, including both public and nonpublic school students. The scores are reported only in individual grade equivalents, so all our analyses use this metric.

---

*The level administered was determined by the student's instructional level rather than grade level.

†We use "cohort" to refer to a group of students who progress through school together.

Our data can be presented in a number of different ways since they include several cohorts and grades as well as both public and nonpublic school students. For simplicity, we present in this section only samples drawn from the larger set of public school students. The corresponding data for the nonpublic school students are contained in Appendix D.

For each grade, we defined two samples of public school students. The first sample consists of students who were tested at least three consecutive times: in fall and in spring of one academic year and in fall of the next year. These students have participated in the reading program for at least the one school year between the fall and spring tests. The second sample is a subset of the first and consists of all students who were tested at least five consecutive times: in fall and in spring of two consecutive academic years and in the following fall. These students have participated in the program for at least the two school years spanned by the test points. Defining the samples in this way allows each mean for a given sample to be based on the same students.

Since we included in our samples only those students with sufficient test points, the samples do not reflect all students in the program for a given year and grade. However, for students tested in a given grade, we have compared each mean from our matched, longitudinal sample to each mean from a larger sample of one-year participants. The means of the one-year participants were consistently higher. However, since the differences were always small and the pattern of gains quite similar, our conclusions were not affected. (See Appendix B for a comparison of the longitudinal and the one-year participant samples.)

In the following tables, we present data by grade for the two samples of public school students (those with three and five consecutive test administrations). These data are presented separately for each cohort, by grade, in Appendix C. With very few exceptions, the patterns described below are consistent across cohorts, grades, and both public and nonpublic school students.*

*We have eliminated means based on fewer than 30 students.

The grade-equivalent means and standard deviations for the samples with three consecutive test points are presented by grade level in Table 1. Hence, for Grade 3 we show the means and standard deviations for the fall and spring of Grade 3 and the fall of Grade 4. The sample sizes for the grades range from 128 to 980. For these samples, Table 2 presents achievement gains and rates in grade-equivalent months based on the means in Table 1. Column I contains the differences between the fall and spring scores, and Column II shows the annual (fall-to-fall) gains. Column III presents the fall-to-spring gains as monthly achievement rates (Column I divided by 7, the average interval between fall and spring tests). Column IV expresses the fall-to-fall gains in terms of monthly rates (Column II divided by 10).

Table 1

CITY M MEANS AND STANDARD DEVIATIONS IN GRADE EQUIVAI NTS
FOR THE GATES-MacGINITIE READING TESTS BY GRADE
FOR PUBLIC SCHOOL STUDENTS WITH AT LEAST THREE CONSECUTIVE TEST POINTS

Grade	N	Fall	Spring	Fall
3	272	2.23 (1.04)	3.29 (1.42)	2.78 (0.96)
4	931	2.65 (0.83)	3.58 (1.19)	3.18 (0.96)
5	980	3.26 (0.99)	4.30 (1.38)	4.01 (1.30)
6	316	3.85 (1.2)	4.78 (1.47)	4.42 (1.32)
7	128	4.35 (1.24)	5.25 (1.68)	4.95 (1.41)

Comparing Columns I and II in Table 2, we see that, in each grade, the achievement gain is substantially lower when based on the fall-to-fall period (Column II) than when based on the traditional fall-to-spring

period (Column I). When the fall-to-spring achievement gains are translated into monthly achievement rates (Column III), all the samples have considerably higher rates than the month-for-month standard, which suggests a highly successful program.

Table 2

CITY M ACHIEVEMENT GAINS AND MONTHLY RATES BASED ON TWO DIFFERENT PERIODS OF TIME FOR PUBLIC SCHOOL STUDENTS WITH AT LEAST THREE CONSECUTIVE TEST POINTS
(Grade-Equivalent Metric)

		Achievement in Grade-Equivalent Months*		Monthly Achievement Rates	
		I	II	III	IV
Grade	N	Fall to Spring	Fall to Fall	Fall to Spring (I ÷ 7)	Fall to Fall (II ÷ 10)
3	272	10.6	5.5	1.5	0.6
4	931	9.3	5.3	1.3	0.5
5	980	10.4	7.5	1.5	0.8
6	316	9.3	5.7	1.3	0.6
7	128	9.0	6.0	1.3	0.6

*The achievement is based on the means in Table 1.

However, the monthly achievement rates in Column IV that are based on the fall-to-fall period present a very different picture. These rates are all much lower than the month-for-month standard. In fact, with one exception, these rates do not reach the seven months expected for disadvantaged students. In the one exception (Grade 5) this expectation is exceeded only slightly. Hence, while the traditional fall-to-spring evaluation periods suggest a highly successful program, the more appropriate fall-to-fall interval suggests an unsuccessful program, one that does not produce sustained achievement.

We now extend our presentation to a two-year evaluation. Tables 3 and 4 display data for the group of students with five consecutive test points. Table 3 presents the grade-equivalent means and standard deviations for the five test points, with sample sizes ranging from 45 to 324.

Table 3

CITY M MEANS AND STANDARD DEVIATIONS IN GRADE EQUIVALENTS FOR THE GATES-MacGINITIE READING TEST BY GRADE FOR PUBLIC SCHOOL STUDENTS WITH AT LEAST FIVE CONSECUTIVE TEST POINTS

Grade	N	I Fall 1	II Spring 1	III Fall 2	IV Spring 2	V Fall 3
3-4	87	2.57 (1.49)	3.62 (1.77)	3.00 (1.10)	4.12 (1.42)	3.53 (0.92)
4-5	324	2.75 (0.94)	3.54 (1.28)	3.11 (0.88)	4.27 (1.24)	3.87 (1.15)
5-6	130	3.22 (1.15)	4.14 (1.47)	3.86 (1.22)	4.81 (1.51)	4.41 (1.25)
6-7	45	4.08 (1.19)	4.91 (1.58)	4.52 (1.07)	5.44 (1.47)	5.04 (1.25)

Table 4 shows the monthly achievement rates for two-year evaluations based on three different time periods (see Figure 1). Columns IA and IB present the rates for each pretest (fall) to posttest (spring) interval. In each grade, the monthly rate for each year equals or exceeds the month-for-month standard. A program that meets or exceeds the month-for-month standard for two consecutive years would undoubtedly be considered quite successful.

Column II presents the monthly achievement rate based on the period from the beginning of the first year (Fall 1) to the end of the second year (Spring 2). This estimate includes the losses incurred during the intervening summer. While the rates are lower than those obtained for each school year separately, they are still close to the month-for-month standard and would likely lead to the conclusion that the program is a success.

Table 4

CITY M TWO-YEAR MONTHLY ACHIEVEMENT RATES BASED ON THREE DIFFERENT PERIODS OF TIME FOR PUBLIC SCHOOL STUDENTS WITH AT LEAST FIVE CONSECUTIVE TEST POINTS*
(Grade-Equivalent Metric)

		I		II	III
		A	B		
		Year 1	Year 2		
Grade	N	(Fall 1 to Spring 1)	(Fall 2 to Spring 2)	Fall 1 to Spring 2	Fall 1 to Fall 3
3-4	87	1.5	1.6	0.9	0.5
4-5	324	1.1	1.7	0.9	0.6
5-6	130	1.3	1.4	0.9	0.6
6-7	45	1.2	1.3	0.8	0.5

*The rates are calculated by dividing the achievement gain for each interval (based on the means in Table 3) by the number of grade-equivalent months in the interval (7 for Columns IA and IB, 17 for Column II, and 20 for Column III).

Column III presents the average monthly achievement rate based on the full two calendar years, from the beginning of the first year (Fall 1) to the beginning of the third year (Fall 3). This estimate includes both the intervening summer and the summer following the second year. We see under this time period that the achievement rates are substantially lower. In all four grades, the two-year rate under this time frame is considerably less than the month-for-month standard. If we compare the average annual achievement† to the expectation of seven months, none of the grades exceeds that expectation. Under this time frame, the program would not be judged a success.

†The average annual achievement in grade-equivalent months for the four grades is 4.7, 5.6, 6.0, and 4.8, respectively. These are calculated by dividing the two-year gain by 2.

While the third time period, Fall 1 to Fall 3, produces a far more discouraging picture than that produced by the other two methods, we believe that this is the appropriate time frame for a two-year evaluation. This period is the only one that captures the extent to which achievement is sustained over the summer months. Since these rates are considerably smaller than those based on periods excluding one or both summer intervals, we conclude that high gains achieved during the school year do not persist until even the beginning of the next school year.

California

Our second source of data was California State Demonstration Programs in Intensive Instruction in Reading and Mathematics. These schoolwide programs are limited to junior high schools located in low-income areas. Participants are predominantly educationally disadvantaged students who meet eligibility requirements for Title I. The programs began in Grade 7 during the 1969-70 academic year, moved with the students to Grade 8 in 1970-71, and to Grade 9 in 1971-72. In 1972-73, the three-year cycle began again. Since the programs are schoolwide, there can be no "graduation" out of the program; all students remain in the program for all three years.

In some districts, other compensatory funds were used to replicate the program in those grades not supported by the state. The data we obtained are from evaluations of both original state demonstration programs and district replications. The data are for one cohort of students in each of three reading programs (one school in each of three cities) and for three cohorts of students in a mathematics program (two cohorts in one school and one cohort in a second school, all in the same city). Hence we have six different samples. For all the programs, the Comprehensive Test of Basic Skills (CTBS) was administered in the fall and spring of Grades 7, 8, and 9, with the reading portion administered in the reading programs and the mathematics portion in the mathematics program. The forms and levels of the CTBS in each sample are shown in Table 5.

Table 5

THE FORMS AND LEVELS OF THE COMPREHENSIVE TEST OF BASIC SKILLS ADMINISTERED TO EACH SAMPLE

	Mathematics, City A			Reading		
	School A Cohort 1	School A Cohort 2	School B	City B	City C	City D
7 Fall	Q3	Q3	Q3	Q3	Q3	Q3
7 Spring	Q3	Q3	Q3	Q3	Q3	Q3
8 Fall	R3	R3	R3	R3	R3	Q3
8 Spring	R3	R3	R3	Q3,R3	R3	Q3
9 Fall	R4	R2,R3,R4	R4	R4	Q4	R4

Each of our six samples contains students who were tested at least the following five times: in fall and in spring of Grades 7 and 8 and in fall of Grade 9. These six samples do not contain all students who participated in the programs since they consist only of students with five test scores. Therefore, we compared the longitudinal samples to larger samples of all two-year participants who were tested at least twice. These comparisons are presented in Appendix E. Appendix F contains the California data separately by grade.

The grade-equivalent means and standard deviations for the samples are presented in Table 6. The sample sizes range from 52 to 153. In Table 7, we consider Grades 7 and 8 separately, and we present for each estimates of achievement rates based on two different time periods. Column I contains estimates of Grade 7 achievement rates based on the traditional pretest (fall) to posttest (spring) interval. For all programs and cohorts, these rates are quite high, ranging from 1.7 to 3.1 months per month. Each rate clearly exceeds the month-for-month standard. Using this rate, all programs would be considered successful.

Column II contains estimates of Grade 7 achievement rates based on the interval from fall of Grade 7 to fall of Grade 8. These rates are all substantially below those found in Column II, ranging from 0.4 to 1.3 months per month. However, three rates still exceed the month-for-month standard. Using the rates in Column II, the reading program in

Table 6

CALIFORNIA DEMONSTRATION PROGRAMS MEANS AND STANDARD DEVIATIONS IN GRADE EQUIVALENTS FOR THE CTBS

	Means				
	7 Fall	7 Spring	8 Fall	8 Spring	9 Fall
Mathematics: City A					
School A					
Cohort 1 (n = 52)	5.4* (1.8)	6.7 (2.2)	6.1 (2.2)	7.7 (2.6)	7.4 (2.6)
Cohort 2 (n = 67)	5.1 (1.6)	7.1 (2.3)	6.2 (2.3)	7.4 (2.4)	7.2 (2.2)
School B (n = 153)	6.0 (1.7)	8.2 (2.2)	7.3 (2.2)	8.7 (2.6)	8.7 (2.3)
Reading					
City B (n = 99)	4.3 (1.7)	5.6 (1.8)	5.5 (1.7)	6.5 (1.7)	6.2 (1.8)
City C (n = 54)	4.6† (1.8)	5.8 (2.0)	5.0 (2.0)	7.5± (2.4)	5.8 (2.4)
City D (n = 47)	5.7 (2.6)	6.9 (2.7)	6.3 (2.7)	7.3 (2.8)	7.8 (3.0)

*The grade-equivalent means were calculated by first computing means in the expanded standard score scale. The standard score means were then converted to associated grade-equivalent scores. The standard deviations were calculated in a similar fashion.

†Since standard scores were not available at this test administration, the means were computed in grade equivalents at the individual level.

±We are suspicious of this high score and have reason to believe that it may reflect unusually permissive testing conditions.

Table 7

CALIFORNIA DEMONSTRATION PROGRAMS MONTHLY ACHIEVEMENT RATES BASED ON TWO DIFFERENT PERIODS OF TIME*

	Grade 7		Grade 8	
	I 7 Fall to 7 Spring	II 7 Fall to 8 Fall	III 8 Fall to 8 Spring	IV 8 Fall to 9 Fall
Mathematics: City A				
School A				
Cohort 1 (n = 52)	1.9	0.7	2.3	1.3
Cohort 2 (n = 67)	2.9	1.1	1.7	1.0
School B (n = 153)	3.1	1.3	2.0	1.4
Reading				
City B (n = 99)	1.9	1.2	1.4	0.7
City C (n = 54)	1.7	0.4	3.6	0.8
City D (n = 47)	1.7	0.6	1.4	1.5

*The rates are calculated by dividing the achievement gain for each interval (based on the means in Table 6) by the number of grade-equivalent months in the interval (7 for Columns I and III, 10 for Columns II and IV).

City B would be still judged successful, as would the mathematics program for Cohort 2 in School A and in School B. The remaining three programs would not be judged successful during the longer, 12-month interval.

Columns III and IV contain estimates of Grade 8 achievement rates based on the traditional fall-to-spring interval and the longer fall-to-fall interval, respectively. As before, the achievement rates based on fall to spring (Column III) are uniformly high, ranging from 1.4 to 3.6 months per month. Using the rates in Column III, all programs would be considered quite successful. As with all our analyses, the rates based on fall to fall (Column IV) are substantially lower, ranging from 0.7 to 1.5 months per month. Using these rates, two of the programs would not be considered successful. However, in four of the programs the lower rates still exceed the 1.0 standard and they would be considered successful.

In Table 8, we present monthly achievement rates for two-year evaluations based on the three different time periods described above (see Figure 1). Columns IA and IB are identical to Columns I and III in Table 7 and give the achievement rates for each pretest-to-posttest interval. These rates are uniformly high, all easily exceeding the month-for-month standard. Programs with such high rates of achievement for two consecutive years would clearly be judged quite successful.

Achievement rates based on the interval from the beginning of Grade 7 to the end of Grade 8 are presented in Column II. These rates are also high, ranging from 0.9 to 1.7 months per month. Though these rates are lower than those calculated for each school year separately, they are sufficiently high for the programs to still be considered successful.

Column III contains the achievement rates based on the 24-month interval from the fall of Grade 7 to the fall of Grade 9. With one exception (City D), for each sample the rate in Column III is lower than both the rate in Column I and the rate in Column II. The achievement rates for five of the six samples equal or exceed the month-for-month

Table 8

CALIFORNIA DEMONSTRATION PROGRAMS TWO-YEAR MONTHLY ACHIEVEMENT RATES BASED ON THREE DIFFERENT PERIODS OF TIME*

	IA Fall to Spring Grade 7	IB Fall to Spring Grade 8	II 7 Fall to 8 Spring	III 7 Fall to 9 Fall
Mathematics: City A				
School A				
Cohort 1 (n = 52)	1.9	2.3	1.4	1.0
Cohort 2 (n = 67)	2.9	1.7	1.4	1.1
School B (n = 153)	3.1	2.0	1.6	1.4
Reading				
City B (n = 99)	1.9	1.4	1.3	1.0
City C (n = 54)	1.7	3.6	1.7	0.6
City D (n = 47)	1.7	1.4	0.9	1.1

*The rates are calculated by dividing the achievement gain for each interval (based on the means in Table 6) by the number of grade-equivalent months in the interval (7 for Columns IA and IB, 17 for Column I, and 20 for Column III).

standard, with one considerably higher (School B, City A). The rate for the sample from City C is substantially below the standard. In spite of the overall lower rates under this time frame, if an evaluation were based on an actual two-year interval, five of the six samples would still lead evaluators to judge the programs as successful.

Again, we argue that the third time period, spanning two calendar years, is the most appropriate for a two-year evaluation. The other time periods do not consider the extent to which achievement is sustained during the summer months. Since some of the achievement is not sustained, the two-year period generally yields lower rates. Yet, for five of the six samples this estimate of sustained achievement is still substantially higher than the expected achievement of 14 months over two years. Hence, five samples represent programs that would be judged successful even under the most demanding time interval.

V CONCLUSIONS

We have analyzed several sets of data, all of which demonstrate that estimates of achievement rates can differ dramatically when different time periods are used for an evaluation of a compensatory education program. Specifically, we find that including the summer months in an evaluation often considerably reduces estimates of achievement and hence monthly achievement rates. In other words, achievement gains made during the school year are not sustained, even until the next fall. These findings are consistently supported by the data regardless of the standardized achievement test used, the grade level, the subject area, or the program. While these data were not selected to be nationally representative of compensatory programs, the consistency of the findings across different tests, grades, subjects, and programs suggests that our conclusions are generalizable.

We conclude that if compensatory education programs are going to be evaluated on the basis of achievement, it is essential to have measures of sustained achievement. We suggest that an evaluation of a one-year program be based at a minimum on fall-to-fall achievement gains. Extending the argument to evaluations of multi-year programs, we suggest that such evaluations be based on a time period that includes the summer interval following each year of the program. By including the summer following the program in an evaluation, the judgment of program success will at least partially reflect the extent to which achievement is sustained and therefore will provide more useful information to those concerned with improving programs. We also urge that a spring test be administered so that school-year and summer achievement can be compared. This can provide a basis for future research on the extent and causes of summer losses, which can in turn lead to appropriate remedies.

Many school districts already administer achievement tests to compensatory education students in both the fall and the spring of each year.

The results of these tests can be used to create longitudinal data, such as we have analyzed in this report. Although the longitudinal data will not be complete for all program participants, those students for whom no data are missing can provide a basis for estimating sustained achievement.* Hence, for these districts we do not recommend any change in their collection of data; rather, we recommend a modification in the way the data are used: In addition to estimates of fall-to-spring achievement, these districts should estimate fall-to-fall achievement.

We also do not recommend any change in the collection of data in those districts that administer achievement tests in the fall only. Although the school-year and summer gains (or losses) cannot be separated, sustained achievement can be estimated by analyzing data from all students with at least two consecutive fall scores. (In this case, biases caused by attrition will be more difficult to investigate. The type of analyses done in Appendix B rely on also having data from achievement tests administered in the spring.)

In the school districts that administer achievement tests only in the spring, we do recommend changes in the collection of data. A spring-to-spring gain consists of two components: the summer gain (or, more probably, loss) made prior to participation in the program and the fall-to-spring gain made while in the program. Therefore, the spring-to-spring gain does not provide any indication of the extent to which the fall-to-spring, school-year gain is sustained. We recommend that these districts also administer tests each fall, so that estimates of sustained gains can be made.

In conclusion, we urge that districts administer achievement tests minimally each fall and preferably each fall and spring. These data would provide the capability for estimating the extent to which school-year gains are sustained through the following summer. Both fall and spring tests have the added advantage of allowing a separation of

---

*Comparisons of the longitudinal sample with other samples will be required to investigate possible biases. See Appendices B and E.

school-year and summer achievement. Although this information is not critical for estimates of annual gains, it is valuable for studying the extent and causes of summer losses. If, for example, the phenomenon is a fuction of the measures used, the standardized achievement tests, one would want to change the measures not the program. If it is a result of instructional techniques that mitigate against retention, then the techniques should be changed. Since there are no simple solutions (for example, there is little research to support the notion that summer school would alleviate the summer losses), it is important to be able to determine why the losses occur in order to develop appropriate remedies.

BIBLIOGRAPHY

Angoff, W. H., "Scales, Norms, and Equivalent Scores," in R. L. Thorndike, ed., Educational Measurement (second edition, American Council on Education, Washington, D. C., 1971).

Cronbach, L. J., Essentials of Psychological Testing (third edition, Harper and Row, New York, 1970).

Harcourt Brace Jovanovich, Development and Use of the Grade Equivalent Scale (The author, New York, 1970).

Tallmadge, G. K. and D. P. Horst, "A Procedural Guide for Validating Achievement Gains in Educational Projects," RMC Report No. UR-240, RMC Research Corporation, Mountain View, California (1975).

Thomas, T. and S. Pelavin, "Patterns in ESEA Title I Reading Achievement," EPRC 4537-12 Stanford Research Institute, Menlo Park, California (March 1976).

U.S. General Accounting Office, "Report to the Congress by the Controller General of the United States, Assessment of Reading Activities Funded Under the Federal Program of Aid for Educationally Deprived Children" (Office of Education, Department of Health, Education, and Welfare, Washington, D. C., 1975).

Appendix A

A COMPARISON OF PATTERNS OF ACHIEVEMENT GAINS IN THREE METRICS

Appendix A

A COMPARISON OF PATTERNS OF ACHIEVEMENT GAINS IN THREE METRICS

The data from the California Demonstration Programs contained Comprehensive Test of Basic Skills (CTBS) raw scores for each student. The raw scores permitted us to perform our analyses in any metric we chose. In the body of the report, we used grade-equivalent scores in order to be consistent with previous studies of the impact of compensatory education on achievement. The purpose of this appendix is to present evidence in support of our contention in Section III that the patterns of interest are not seriously affected by the metric.

Table A-1 presents the grade-equivalent means, the standard score means, and the percentile associated with each standard score mean for samples of students who were tested in the fall and spring of Grade 7 and in fall of Grade 8. For each of the samples, the three metrics show the same pattern: the mean substantially increases from fall to spring of Grade 7 and decreases from spring of Grade 7 to the fall of Grade 8. The similarity in patterns is also reflected in Columns I and II of Table A-2. Columns I and II, respectively, present the mean fall-to-spring and fall-to-fall achievement gains. No matter which metric is used, the fall-to-spring mean gain is larger than the fall-to-fall gain. The last column in Table A-2 presents the ratio of the fall-to-fall gain to the fall-to-spring gain. This ratio is the proportion of the fall-to-spring gain that is sustained during the summer. (If the entire gain were sustained, the ratio would be 1.) All three metrics demonstrate the same finding: that a large portion of the fall-to-spring gain is not sustained over the summer. This finding is most pronounced when the ratio is expressed in percentiles. The ratios of gains in grade equivalents and standard scores are quite similar.

Table A-1

GRADE EQUIVALENTS AND STANDARD SCORE MEANS AND THE PERCENTILES ASSOCIATED WITH THE STANDARD SCORES FOR THE CTBS FOR STUDENTS TESTED 7 FALL, 7 SPRING, AND 8 FALL*

		7 Fall	7 Spring	8 Fall
Mathematics				
City A				
School A				
Cohort 1	GE	5.1	6.4	5.8
(n = 109)	SS	413	459	439
Percentile		14	26	15
Cohort 2	GE	5.3	7.3	6.4
(n = 82)	SS	422	495	463
Percentile		18	42	23
School B	GE	5.9	8.1	7.1
(n = 241)	SS	445	523	489
Percentile		27	56	34
Reading				
City B	GE	4.5	5.7	5.6
(n = 102)	SS	405	453	446
Percentile		11	19	15
City C	GE	4.6	5.8	5.1
(n = 65)	SS	415*	456	427
Percentile		14	20	10
City D	GE	5.5	6.5	6.0
(n = 81)	SS	444	478	459
Percentile		23	28	19

*Since standard scores for individuals were not available, the means was calculated in grade equivalents and translated into a standard score.

Table A-2

GRADE 7 ACHIEVEMENT GAINS IN GRADE EQUIVALENTS, STANDARD SCORES, AND PERCENTILES, AND THE RATIO OF THESE GAINS FOR STUDENTS TESTED 7 FALL, 7 SPRING, AND 8 FALL

	Achievement Gain		III
	I Fall to Spring	II Fall to Fall	Ratio of II to I (II ÷ I)
Math			
City A			
School A			
Cohort 1 GE	1.3	0.7	0.54
(n = 109) SS	46	26	0.57
Percentile	12	1	0.08
Cohort 2 GE	2.0	1.1	0.55
(n = 82) SS	73	41	0.56
Percentile	24	5	0.21
School B GE	2.2	1.2	0.55
(n = 241) SS	78	44	0.56
Percentile	29	7	0.24
Reading			
City B GE	1.2	1.1	0.92
(n = 102) SS	48	41	0.85
Percentile	8	4	0.05
City C GE	1.2	0.5	0.42
(n = 65) SS	41	12	0.29
Percentile	6	-4	-0.67
City D GE	1.0	0.5	0.5
(n = 81) SS	34	15	0.44
Percentile	5	-4	-0.80

Appendix B

THE REPRESENTATIVENESS OF THE CITY M SAMPLES

Appendix B

THE REPRESENTATIVENESS OF THE CITY M SAMPLES

The purpose of this appendix is to determine if the City M samples presented in the text were representative of all students who had been in the program for at least one year. To determine this, we formed samples of one-year participants consisting of all students tested in both fall and spring but not tested the following fall. Hence these samples contain all tested students who were definitely in the program one year, but who were not included in the longitudinal samples. We then compared the fall and spring means of the matched, longitudinal samples from the text to the corresponding means of the samples of one-year participants. These comparisons are shown in Columns I and II of Table B-1. The matched, longitudinal samples are labeled M-L and the samples of one-year participants are labeled 1-Y.

In all grades, the means for the one-year participants are higher than those for the longitudinal samples. However, we are primarily concerned with these differences only to the extent that they suggest that the patterns of achievement (school-year and summer) are different for the two samples. For all grades except Grade 3, the means are equally higher in both fall and spring, hence the school-year (fall-to-spring) gains are quite similar for the two samples (see Column IV). We cannot test directly the summer achievement for the two samples since the one-year participants do not have the second fall test. Since the school-year gains are similar, we would expect summer gains (losses) to be similar. But it is conceivable that students with higher mean scores, such as the one-year participants, might show a different achievement pattern over the summer.

To investigate the possibility that there is a difference in summer losses, we formed subsamples of each matched, longitudinal sample. These subsamples consisted of only the highest scoring 60 percent to 80 percent

Table B-1

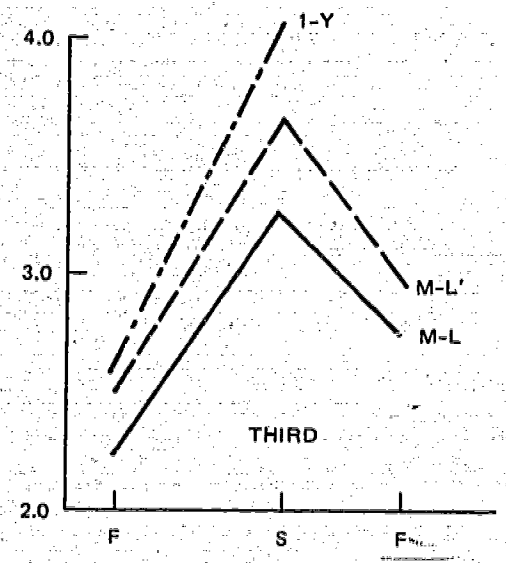
MEANS AND ACHIEVEMENT GAINS FOR THREE SAMPLES

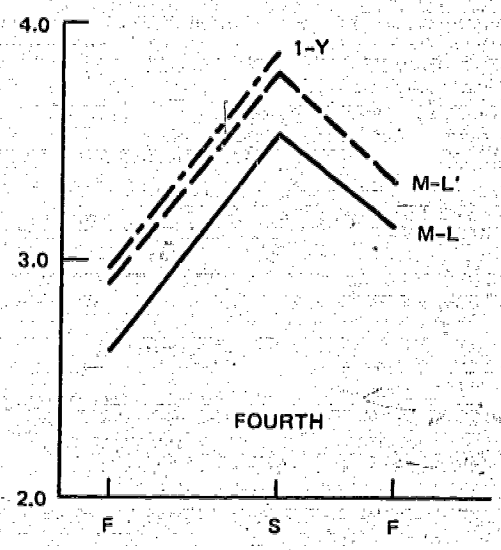
		Means			Achievement Gains	
		I	II	III	IV	V
Grade	N	Fall	Spring	Fall	Fall to Spring	Fall to Fall
3						
M-L	272	2.23	3.29	2.78	1.06	0.55
1-Y	76	2.63	4.07		1.44	
M-L′	177	2.58	3.68	2.97	1.10	0.39
4						
M-L	931	2.65	3.58	3.18	0.93	0.55
1-Y	496	2.93	3.91		0.98	
M-L′	705	2.91	3.82	3.35	0.91	0.44
5						
M-L	980	3.26	4.30	4.01	1.04	0.75
1-Y	582	3.53	4.51		0.98	
M-L′	801	3.53	4.55	4.26	1.02	0.73
6						
M-L	316	3.85	4.78	4.42	0.93	0.57
1-Y	1141	4.28	5.17		0.89	
M-L′	245	4.27	5.09	4.70	0.82	0.43
7						
M-L	128	4.35	5.25	4.95	0.90	0.60
1-Y	318	4.63	5.54		0.91	
M-L′	112	4.61	5.48	5.18	0.87	0.57

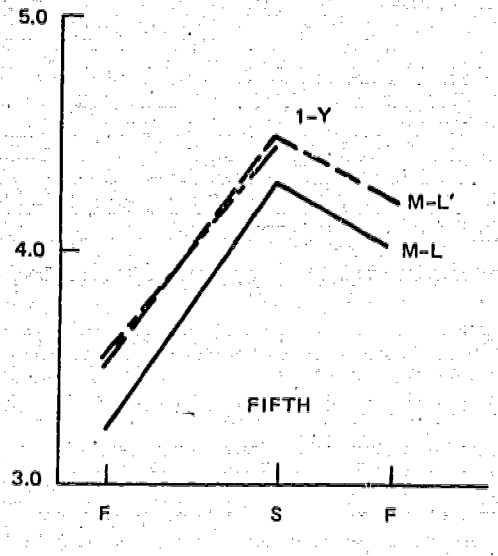
of each matched, longitudinal sample, based on the first fall scores. By including only the higher scoring students, the initial fall means of the subsamples were all within 0.05 grade-equivalent years of the fall mean scores of the 1-Y samples. Given similar fall scores, we assume that if the subsamples have the same school-year gains, they will have similar summer losses. Hence we first compare their achievement gains during the school year. The means and achievement gains for the subsamples, labeled M-L′, appear in Table B-1. A comparison of 1-Y with M-L′ shows that in all cases, with the exception of Grade 3, the difference between the spring mean scores is less than 0.1 grade-equivalent years. Therefore, during the school year the two groups show exactly the same pattern of achievement gains.

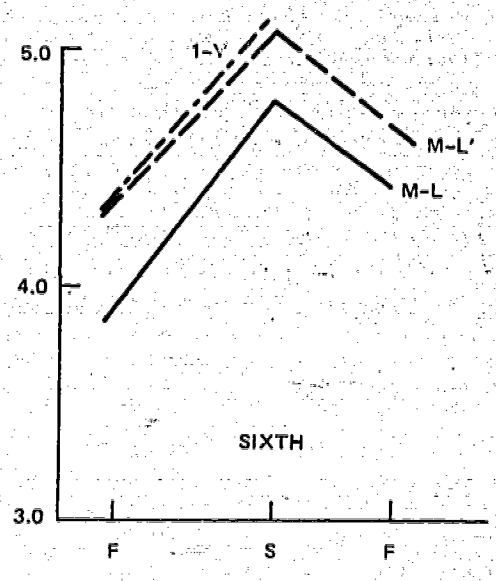
If students who score higher in the first fall and spring have a different achievement pattern over the summer, then it should be apparent in the M-L′ subsamples. We therefore compare the school-year and summer gains of the higher scoring M-L′ samples to the M-L samples. From Column IV, Table B-1, it is evident that the school-year gains and summer losses are remarkably similar for the two sets of samples. This similarity is shown graphically in Figure B-1. Figure B-1 presents the gains by grade for each of the three samples described above. Each chart demonstrates the similarity of the fall test point for the one-year participants (1-Y) and the higher scoring, longitudinal subsamples (M-L′). With the exception of Grade 3, the school-year gains for these two samples are also quite similar. Each chart also demonstrates the similarity between both the school-year gains and summer losses for the entire matched, longitudinal samples (M-L) and the higher scoring subsamples (M-L′).

In summary, our first analysis demonstrates that the school-year gains are quite similar for the matched, longitudinal samples (M-L) and the one-year participants excluded from these samples (1-Y). The second analysis investigates whether this similarity also holds for the summer losses. We found that the higher scoring students in the longitudinal samples (M-L′), selected to match the fall scores of the one-year participants samples (1-Y), have spring scores and hence school-year gains virtually identical to the one-year participants. We also demonstrated









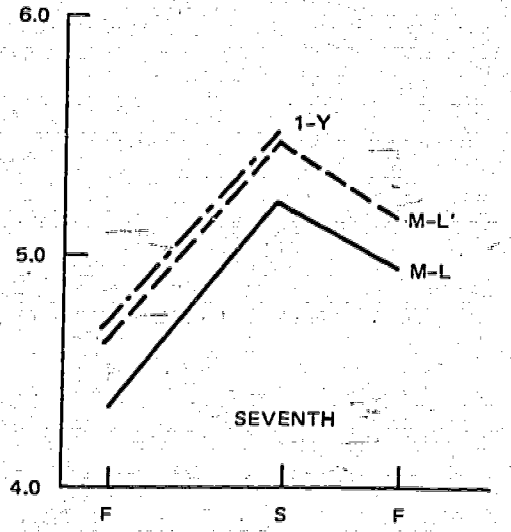


FIGURE B-1
PATTERNS OF ACHIEVEMENT
FOR THREE SAMPLES

that the summer losses for the higher scoring students in the longitudinal samples (M-L') are the same as those for the entire matched, longitudinal samples (M-L). We conclude that if we had the second fall score for the one-year participants not included in the longitudinal samples, they would have demonstrated the same pattern of summer losses.

Appendix C

PUBLIC SCHOOL STUDENTS IN CITY M BY COHORT

Appendix C

PUBLIC SCHOOL STUDENTS IN CITY M BY COHORT

Appendix C contains additional information on the public school students in City M. Table C-1 describes the cohorts for which we have data by year and grade level. Each cohort represents a wave of students and is assigned the number that was the student's grade in 1971-72. Thus, we have two years of data for Cohort 1, three years of data for Cohort 2, and four years of data for Cohorts 3-6.

Table C-2 presents the means and standard deviations for the samples defined by public school students with at least three consecutive test points for each cohort. The totals, which are averaged across cohorts, are the figures presented in Table 1 in the text. Table C-3 corresponds to Table 2 in the text and presents the gains and rates over two time periods, separately for each cohort.

Tables C-4 and C-5 are analagous to C-2 and C-3 but are based on the sample defined by public school students with at least five consecutive test points. The totals in Table C-4 correspond to the figures in Table 3 in the text. Table C-5 presents the achievement rates based on three time periods for each cohort, and corresponds to Table 4 in the text.

Table C-1

NAMES OF COHORTS BY YEAR BY GRADE

Grade	1971-72	1972-73	1973-74	1974-75
3	Cohort 3	Cohort 2	Cohort 1	
4	Cohort 4	Cohort 3	Cohort 2	Cohort 1
5	Cohort 5	Cohort 4	Cohort 3	Cohort 2
6	Cohort 6	Cohort 5	Cohort 4	Cohort 3
7		Cohort 6	Cohort 5	Cohort 4
8			Cohort 6	Cohort 5
9				Cohort 6

Table C-2

MEANS AND STANDARD DEVIATIONS IN GRADE EQUIVALENTS
FOR THE GATES-MacGINITIE READING TEST BY COHORT FOR
PUBLIC SCHOOL STUDENTS WITH AT LEAST THREE CONSECUTIVE TEST POINTS

		N	Fall	Spring	Fall
Grade 3 -	Cohort 1	152	1.97 (0.45)	3.05 (1.10)	2.60 (0.78)
	Cohort 2	90	2.67 (1.57)	3.72 (1.86)	2.82 (0.64)
	Cohort 3	30	2.20 (0.71)	3.26 (0.96)	3.54 (1.84)
	Totals	272	2.23 (1.04)	3.29 (1.42)	2.78 (0.96)
Grade 4 -	Cohort 2	387	2.54 (0.65)	3.60 (1.04)	3.16 (0.88)
	Cohort 3	270	2.81 (1.13)	3.63 (1.48)	3.07 (0.88)
	Cohort 4	274	2.64 (0.67)	3.51 (1.07)	3.31 (1.13)
	Totals	931	2.65 (0.83)	3.58 (1.19)	3.18 (0.96)
Grade 5 -	Cohort 3	401	3.25 (0.90)	4.45 (1.30)	4.10 (1.32)
	Cohort 4	278	3.46 (1.22)	4.52 (1.62)	3.99 (1.22)
	Cohort 5	301	3.09 (0.83)	3.89 (1.14)	3.92 (1.34)
	Totals	980	3.26 (0.99)	4.30 (1.38)	4.01 (1.30)
Grade 6 -	Cohort 4	92	3.59 (0.99)	4.81 (1.36)	4.41 (1.22)
	Cohort 5	136	4.08 (1.38)	4.79 (1.64)	4.45 (1.31)
	Cohort 6	88	3.78 (1.02)	4.75 (1.29)	4.38 (1.43)
	Totals	316	3.85 (1.20)	4.78 (1.47)	4.42 (1.32)
Grade 7 -	Cohort 5	73	4.37 (1.19)	5.30 (1.58)	4.84 (1.34)
	Cohort 6	55	4.33 (1.32)	5.18 (1.82)	5.10 (1.50)
	Totals	130	4.35 (1.24)	5.25 (1.68)	4.95 (1.41)

Table C-3

ACHIEVEMENT GAINS AND MONTHLY RATES BASED ON TWO DIFFERENT PERIODS OF TIME FOR PUBLIC SCHOOL STUDENTS WITH AT LEAST THREE CONSECUTIVE TEST POINTS

		Achievement in Grade-Equivalent Months*		Monthly Achievement Rates	
		I	II	III	IV
Grade	N	Fall to Spring	Fall to Fall	Fall to Spring (I ÷ 7)	Fall to Fall (II ÷ 10)
Grade 3					
Cohort 1	152	10.8	6.3	1.5	0.6
Cohort 2	90	10.5	1.5	1.5	0.2
Cohort 3	30	10.6	13.4	1.5	1.3
Total	27[illegible]	10.6	5.5	1.5	0.6
Grade 4					
Cohort 2	387	10.6	6.2	1.5	0.6
Cohort 3	270	8.3	2.6	1.2	0.3
Cohort 4	274	8.7	6.7	1.2	0.7
Total	931	9.3	5.3	1.3	0.5
Grade 5					
Cohort 3	401	12.0	8.5	1.7	0.8
Cohort 4	278	10.6	5.3	1.5	0.5
Cohort 5	301	8.0	8.3	1.1	0.8
Total	980	10.4	7.5	1.5	0.8
Grade 6					
Cohort 4	92	12.2	8.2	1.7	0.8
Cohort 5	136	7.1	3.7	1.0	0.4
Cohort 6	88	9.7	6.0	1.4	0.6
Total	316	9.3	5.7	1.3	0.6
Grade 7					
Cohort 5	73	9.3	4.7	1.3	0.5
Cohort 6	55	8.5	7.8	1.2	0.8
Total	128	9.0	6.0	1.3	0.6

*Achievement is based on the means in Table C-2

Table C-4

MEANS AND STANDARD DEVIATIONS IN GRADE EQUIVALENTS
FOR THE GATES-MacGINITIE READING TEST BY COHORT
FOR PUBLIC SCHOOL STUDENTS WITH AT LEAST FIVE CONSECUTIVE TEST POINTS

	N	I Fall 1	II Spring 1	III Fall 2	IV Spring 2	V Fall 3
Grade 3-4						
Cohort 2	65	2.73 (1.67)	3.80 (1.96)	2.88 (0.67)	3.99 (1.11)	3.51 (0.90)
Cohort 3	22	2.12 (0.60)	3.10 (0.82)	3.38 (1.84)	4.51 (2.07)	3.59 (0.99)
Totals	87	2.57 (1.49)	3.62 (1.77)	3.00 (1.10)	4.12 (1.42)	3.53 (0.92)
Grade 4-5						
Cohort 3	178	2.83 (1.13)	3.65 (1.49)	3.04 (0.84)	4.29 (1.28)	3.85 (1.17)
Cohort 4	146	2.66 (0.63)	3.41 (0.96)	3.19 (0.92)	4.24 (1.19)	3.89 (1.12)
Totals	324	2.75 (0.94)	3.54 (1.28)	3.11 (0.88)	4.27 (1.24)	3.87 (1.15)
Grade 5-6						
Cohort 4	50	3.39 (1.43)	4.33 (1.86)	3.57 (1.01)	4.89 (1.40)	4.29 (1.16)
Cohort 5	80	3.11 (0.92)	4.02 (1.16)	4.04 (1.32)	4.76 (1.58)	4.48 (1.30)
Totals	130	3.22 (1.15)	4.14 (1.47)	3.86 (1.22)	4.81 (1.51)	4.41 (1.25)
Grade 6-7						
Cohort 5	45	4.08 (1.19)	4.91 (1.58)	4.52 (1.07)	5.44 (1.47)	5.04 (1.25)
Cohort 6	18	3.76 (1.04)	4.17 (0.99)	4.16 (1.46)	5.19 (1.99)	4.98 (1.08)
Totals	63	3.99 (1.14)	4.70 (1.45)	4.42 (1.17)	5.37 (1.61)	5.02 (1.19)

Table C-5

TWO-YEAR MONTHLY ACHIEVEMENT RATES
BASED ON THREE DIFFERENT PERIODS OF TIME
FOR PUBLIC SCHOOL STUDENTS WITH AT LEAST FIVE CONSECUTIVE TEST POINTS

		I		II	III
		A	B		
		Year 1	Year 2		
Grade	N	Fall 1 to Spring 1	Fall 2 to Fall 2	Fall 1 to Spring 2	Fall 1 to Fall 3
Grade 3-4					
Cohort 2	65	1.5 (10.7)	1.6 (11.1)	0.7 (12.6)	0.4 (7.8)
Cohort 3	22	1.4 (9.8)	1.6 (11.3)	1.4 (23.9)	0.7 (14.7)
Total	87	1.5 (10.5)	1.6 (11.2)	0.9 (15.5)	0.5 (9.6)
Grade 4-5					
Cohort 3	178	1.2 (8.2)	1.8 (12.5)	0.9 (14.6)	0.5 (10.2)
Cohort 4	146	1.1 (7.5)	1.5 (10.5)	0.9 (15.8)	0.6 (12.3)
Total	324	1.1 (7.9)	1.7 (11.6)	0.9 (14.6)	0.6 (11.2)
Grade 5-6					
Cohort 4	50	1.3 (9.4)	1.9 (13.2)	0.9 (15.0)	0.5 (9.0)
Cohort 5	80	1.3 (9.1)	1.0 (7.2)	1.0 (16.5)	0.7 (13.7)
Total	130	1.3 (9.2)	1.3 (9.5)	0.9 (15.9)	0.6 (11.9)
Grade 6-7					
Cohort 5	45	1.2 (8.3)	1.3 (9.2)	0.8 (13.6)	0.5 (9.6)
Cohort 6	18	0.6 (4.1)	1.5 (10.3)	0.8 (14.3)	0.6 (12.2)
Total	63	1.0 (7.2)	1.3 (9.5)	0.8 (13.8)	0.5 (10.3)

Appendix D

STUDENTS IN CITY M NONPUBLIC SCHOOLS

Appendix D

STUDENTS IN CITY M NONPUBLIC SCHOOLS

Appendix D contains the data for the nonpublic school students in City M. These tables correspond to those presented in the text and in Appendix C for the public school students. The cohort names are the same as those presented in Appendix C, Table C-1. The reading program was not begun in the nonpublic schools until 1972-73. Therefore, Tables D-1 through D-6 span fewer grades than the corresponding tables for the public school students. Table D-1 presents the means and standard deviations (for the nonpublic school students with at least three consecutive test points) for each cohort. The achievement gains and rates over two time periods for these students are given in Table D-2. Tables D-3 and D-4 contain the same data as D-1 and D-2, respectively, averaged over the cohorts.

Tables D-5 and D-6 present data for the non-public school students with at least five consecutive test administrations. Table D-5 contains the means and standard deviations for these students and Table D-6 presents monthly achievement based on three time periods.

Table D-1

MEANS AND STANDARD DEVIATIONS IN GRADE EQUIVALENTS FOR THE GATES-MacGINITIE READING TEST BY COHORT FOR NONPUBLIC SCHOOL STUDENTS WITH AT LEAST THREE CONSECUTIVE TEST POINTS

	N	Fall	Spring	Fall
Grade 4				
Cohort 2	69	3.34 (0.84)	4.40 (1.40)	4.09 (1.29)
Cohort 3	78	2.95 (0.76)	3.75 (1.00)	3.77 (1.12)
Grade 5				
Cohort 3	89	3.66 (0.99)	5.01 (1.14)	4.54 (1.13)
Cohort 4	88	3.66 (1.01)	4.75 (1.44)	4.74 (1.36)
Grade 6				
Cohort 4	49	4.33 (1.08)	5.13 (1.28)	4.76 (1.34)
Cohort 5	68	4.43 (1.30)	5.14 (1.53)	5.19 (1.50)
Grade 7				
Cohort 5	38	4.73 (1.28)	5.68 (1.88)	5.60 (1.61)
Cohort 6	36	5.21 (1.46)	5.61 (2.21)	5.68 (1.52)

Table D-2

ACHIEVEMENT GAINS AND MONTHLY RATES BY COHORT BASED ON TWO DIFFERENT PERIODS OF TIME FOR NONPUBLIC SCHOOL STUDENTS WITH AT LEAST THREE CONSECUTIVE TEST POINTS

	N	Achievement in Grade-Equivalent Months*		Monthly Achievement Rates	
		I	II	III	IV
		Fall to Spring	Fall to Fall	Fall to Spring (I ÷ 7)	Fall to Fall (II ÷ 10)
Grade 4					
Cohort 2	69	10.6	7.5	1.5	0.8
Cohort 3	78	8.0	8.2	1.1	0.8
Grade 5					
Cohort 3	89	13.5	8.8	1.9	0.9
Cohort 4	88	10.9	10.8	1.6	1.1
Grade 6					
Cohort 4	49	8.0	4.3	1.1	0.4
Cohort 5	68	7.1	7.6	1.0	0.8
Grade 7					
Cohort 5	38	9.5	8.7	1.4	0.9
Cohort 6	36	4.0	4.7	0.6	0.5

*The achievement is based on the means in Table D-1.

Table D-3

MEANS AND STANDARD DEVIATIONS IN GRADE EQUIVALENTS FOR THE GATES-MacGINITIE READING TEST BY GRADE FOR NONPUBLIC SCHOOL STUDENTS WITH AT LEAST THREE CONSECUTIVE TEST POINTS

Grade	N	I Fall	II Spring	III Fall
4	147	3.13 (0.82)	4.06 (1.24)	3.92 (1.21)
5	177	3.66 (1.00)	4.88 (1.30)	4.64 (1.25)
6	117	4.39 (1.21)	5.14 (1.42)	5.01 (1.44)
7	74	4.96 (1.38)	5.65 (2.05)	5.68 (1.56)

Table D-4

ACHIEVEMENT GAINS AND MONTHLY RATES BY GRADE BASED ON TWO DIFFERENT PERIODS OF TIME FOR NONPUBLIC SCHOOL STUDENTS WITH AT LEAST THREE CONSECUTIVE TEST POINTS

Grade	N	Achievement in Grade Equivalent-Months* Fall to Spring	Fall to Fall	Monthly Achievement Rates Fall to Spring (I ÷ 7)	Fall to Fall (II ÷ 10)
4	147	9.3	7.9	1.3	0.8
5	177	12.2	9.8	1.7	1.0
6	117	7.5	6.2	1.1	0.6
7	74	6.9	6.8	1.0	0.7

*The achievement is based on the means in Table D-3.

Table D-5

MEANS AND STANDARD DEVIATIONS IN GRADE EQUIVALENTS FOR THE GATES-MacGINITIE READING TEST BY COHORT FOR NONPUBLIC SCHOOL STUDENTS WITH AT LEAST FIVE CONSECUTIVE TEST POINTS

	N	Fall 1	Spring 1	Fall 2	Spring 2	Fall 3
Grade 4-5						
Cohort 3	59	2.84 (0.75)	3.62 (0.93)	3.61 (0.99)	4.96 (1.16)	4.45 (1.12)
Grade 5-6						
Cohort 4	39	3.38 (0.98)	4.49 (1.29)	4.43 (1.08)	5.14 (1.22)	4.74 (1.24)
Grade 6-7						
Cohort 5	28	4.36 (1.35)	4.88 (1.31)	4.85 (1.31)	5.67 (2.08)	5.75 (1.44)

Table D-6

TWO-YEAR MONTHLY ACHIEVEMENT BASED ON THREE DIFFERENT TIME PERIODS FOR NONPUBLIC SCHOOL STUDENTS WITH AT LEAST FIVE CONSECUTIVE TEST POINTS*

		I		II	III
		A Year 1	B Year 2		
	N	Fall 1 to Spring 1	Fall 2 to Spring 2	Fall 1 to Spring 2	Fall 1 to Fall 3
Grade 4-5					
Cohort 3	59	1.1 (7.8)	1.9 (13.5)	1.2 (21.2)	0.8 (16.1)
Grade 5-6					
Cohort 4	39	1.6 (11.1)	1.0 (7.1)	1.0 (17.6)	0.7 (13.6)
Grade 6-7					
Cohort 5	28	0.7 (5.2)	1.2 (8.2)	0.8 (13.1)	0.7 (13.9)

*The achievement in months is in parentheses. This achievement is based on the means in Table D-5.

Appendix E

THE REPRESENTATIVENESS OF THE CALIFORNIA SAMPLES

Appendix E

THE REPRESENTATIVENESS OF THE CALIFORNIA SAMPLES

The purpose of this appendix is to determine if the California samples presented in the text are representative of all students who had participated in the programs during both Grade 7 and Grade 8. To form samples of two-year participants, we included all students who were tested in the fall of Grade 7 and in the spring of Grade 8, the fall of Grade 9, or the spring of Grade 9. We then compare the fall and spring means of the samples of two-year participants to the means of the matched, longitudinal sample from the text. This comparison is shown in Table E-1. The samples from the text are labeled M-L and the sample of two-year participants are labeled 2-Y. City C is not included because we did not have access to the necessary data. Of the five samples presented, the five means are almost identical for two of the samples: City A, School B and City B. Therefore we conclude that the longitudinal samples for these two groups are representative of the larger samples. For Cohort 2 of City A, School A, the first four means are virtually identical for the two samples, but the Grade 9 fall mean is lower for two-year participants (sample 2-Y). Since this results in an even more precipitous summer drop, the difference does not affect our conclusions.

For Cohort 1 of City A, School A, and City D, the means for the two samples are quite different. For Cohort 1, the first four means are consistently lower for the two-year participants, indicating that the longitudinal sample is biased. The patterns of achievement are still similar for the two groups: large gains over the school year and a large loss over the summer. However, the direction of the difference is reversed for the Grade 9 fall test point. Since the 2-Y sample size is reduced considerably between Grade 8 spring and Grade 9 fall, this may not be surprising. For City D, all five means are consistently lower for participants (a range of four to six months). Since these differences are

consistent across all five test points, the patterns of interest are not affected.

Overall, we conclude that the longitudinal samples are representative of the larger sample with the possible exception of Cohort 1 in the fall of Grade 9. Although the means differ somewhat, the pattern of gains over the school year and losses over the summer are similar for the larger samples.

Table E-1

A COMPARISON OF MEANS FOR TWO SAMPLES

		7 Fall	7 Spring	8 Fall	8 Spring	9 Fall	N
Math							
City A							
School A							
Cohort 1 -	M-L	5.4	6.7	6.1	7.7	7.4	52
	2-Y	5.0	6.4	5.6	7.0	7.5	
		n=105	n=99	n=93	n=97	n=67	
Cohort 2 -	M-L	5.1	7.1	6.2	7.4	7.2	67
	2-Y	5.1	7.0	6.2	7.5	6.7	
		n=102	n=95	n=84	n=85	n=86	
School B	M-L	6.0	8.2	7.3	8.7	8.7	153
	2-Y	5.9	8.1	7.1	8.7	8.7	
		n=217	n=206	n=214	n=208	n=167	
Reading							
City B	M-L	4.3	5.6	5.5	6.5	6.2	99
	2-Y	4.3	5.6	5.4	6.5	6.2	
		n=116	n=105	n=111	n=112	n=114	
City C							
City D	M-L	5.7	6.9	6.3	7.3	7.8	47
	2-Y	5.2	6.3	5.7	6.7	7.4	
		n=89	n=85	n=82	n=82	n=55	

Appendix F

CALIFORNIA STUDENTS TESTED IN THE FALL, SPRING, AND FALL

Appendix F

CALIFORNIA STUDENTS TESTED IN THE FALL, SPRING, AND FALL

Tables F-1 through F-4 present additional data in grade-equivalents for the California Demonstration Projects. Table F-1 presents the means and standard deviations for students who were tested in the fall and spring of Grade 7 and the fall of Grade 8. Table F-2 gives the achievement gains and rates for each sample based on the means in Table F-1. Tables F-3 and F-4 correspond exactly to Tables F-1 and F-2 and are based on the students who were tested in the fall and spring of Grade 8 and the fall of Grade 9.

Table F-1

MEANS AND STANDARD DEVIATIONS IN GRADE EQUIVALENTS FOR THE CTBS FOR STUDENTS TESTED 7 FALL, 7 SPRING, AND 8 FALL

	N	7 Fall	7 Spring	8 Fall
Mathematics				
City A				
School A				
Cohort 1	109	5.0 (1.7)	6.2 (2.2)	5.6 (2.2)
Cohort 2	82	5.2 (1.6)	7.3 (2.2)	6.3 (2.2)
School B	241	5.8 (1.6)	8.1 (2.0)	7.1 (2.0)
Reading				
City B	102	4.3 (1.7)	5.6 (1.8)	5.4 (1.8)
City C	65	4.6* (1.7)	5.7 (2.0)	4.9 (1.9)
City D	81	5.4 (2.2)	6.4 (2.6)	5.8 (2.6)

*Since Standard Scores were not available at this test administration, the mean was computed in grade-equivalents at the individual level.

Table F-2

GRADE 7 ACHIEVEMENT GAINS AND MONTHLY RATES BASED ON TWO DIFFERENT PERIODS OF TIME FOR STUDENTS TESTED 7 FALL, 7 SPRING, AND 8 FALL

	N	Achievement in Grade-Equivalent Months* I Fall to Spring	II Fall to Fall	Monthly Achievement Rates III Fall to Spring (I ÷ 7)	IV Fall to Fall (II ÷ 10)
Mathematics					
City A					
School A					
Cohort 1	109	12	6	1.7	0.6
Cohort 2	82	21	11	3.0	1.1
School B	241	23	13	3.3	1.3
Reading					
City B	102	13	11	1.9	1.1
City C	65	11	3	1.6	0.3
City D	81	10	4	1.4	0.4

*Based on means in Table F-1.

Table F-3

MEANS AND STANDARD DEVIATIONS IN GRADE EQUIVALENTS FOR THE CTBS FOR STUDENTS TESTED 8 FALL, 8 SPRING, AND 9 FALL

	N	8 Fall	8 Spring	9 Fall
Mathematics				
City A				
School A				
Cohort 1	68	5.8 (2.2)	7.4 (2.6)	7.3 (2.4)
Cohort 2	89	6.0 (2.2)	7.2 (2.4)	6.9 (2.2)
School B	187	7.1 (2.2)	8.6 (2.4)	8.6 (2.2)
Reading				
City B	107	5.5 (1.7)	6.5 (1.8)	6.2 (1.8)
City C	64	5.0 (1.9)	7.5* (2.4)	5.8 (2.4)
City D	64	6.2 (2.6)	7.1 (2.8)	7.6 (2.8)

*We are suspicious of this high score and have reason to believe that it may reflect unusually permissive testing conditions.

Table F-4

GRADE 8 ACHIEVEMENT GAINS AND MONTHLY RATES BASED ON TWO DIFFERENT PERIODS OF TIME FOR STUDENTS TESTED
8 FALL, 8 SPRING, AND 9 FALL

	N	Achievement in Grade-Equivalent Months* I Fall to Spring	II Fall to Fall	Monthly Achievement Rates III Fall to Spring (I ÷ 7)	IV Fall to Fall (II ÷ 10)
Mathematics					
City A					
School A					
Cohort 1	68	16	15	2.3	1.5
Cohort 2	89	12	9	1.7	0.9
School B	187	15	15	2.1	1.5
Reading					
City B	107	10	7	1.4	0.7
City C	64	25	8	3.6	0.8
City D	64	9	14	1.3	1.4

* Based on means in Table F-3