DOCUMENT RESUME

ED 074 059 SP 006 286

TITLE	Achieving Academic and Social Objectives in Elementary Grades Through Behavioral Analysis. Volume VI, No. 4.
INSTITUTION	Atlanta Public Schools, Ga.; Georgia State Univ., Atlanta. Center for Applied Behavioral Research.
PUB DATE	Jul 72
NOTE	86p.
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Academic Achievement; *Behavioral Objectives; *Behavioral Science Research; Behavior Chaining; *Behavior Change; Elementary School Students; *Social Behavior

ABSTRACT

A Behavior Modification Program was implemented in four classes of a normal public school. These four classes encompassed 93 children ranging in age from 7 to 14 years. Included in the project was a second grade, a third grade, a sixth grade, and a seventh grade. The primary focus of the program in the second, sixth, and seventh grades was upon increasing the academic performance of the children through a variety of methods, including the use of student tutors and points for correct work. In the third grade the primary focus was upon control of extremely high rates of disorderly behavior through the use of a home-based management system. The results revealed that in each of the three classes where academic performance was the primary variable, statistically significant differences were obtained on the Metropolitan Achievement Tests when compared to control groups. Further, all classes increased in number of academic levels passed. Additional experiments indicated that a) it was the point reinforcement system which controlled the children's rate of academic performance, b) student tutors led to increased levels and academic performance when compared to periods when no tutors were present, and c) the academic performance of the tutors was facilitated by the process of teaching other children. A two-item bibliography is included. (Author)

R E S E A R C H A N D D E V E L O P M E N T R E P O R T

Vol. VI, No. 4 July, 1972

ACHIEVING ACADEMIC AND SOCIAL OBJECTIVES IN ELEMENTARY GRADES THROUGH BEHAVIORAL ANALYSIS

Prepared by

The Center for Applied Behavior Research
Georgia State University

Teodoro Ayllon, Ph.D., Director

Assisted by

F. Gerald McCullen, M.A.

Dr. Jarvis Barnes
Assistant Superintendent
for Research and Development

Dr. John W. Letson
Superintendent

Atlanta Public Schools
224 Central Avenue, S. W.
Atlanta, Georgia

PREFACE

Since maintaining orderly behavior within the classroom has traditionally been a major problem for teachers, many studies have been conducted focusing on this area of concern. Research in the area of applied behavior analysis has demonstrated that procedures based on systematic principles of reinforcement can be effective in maintaining classroom discipline; therefore, studies were made in the Atlanta Public School System to determine whether a minimum of disruption should give improved academic performance. The results have indicated that there was no facilitation of academic achievement despite the elimination of classroom disruption, however, it was found that focusing on academic performance, the disruptive behavior was drastically reduced.

The Achieving Academic and Social Objectives in Elementary Grades Through Behavioral Analysis is the result of a three-year study in the Atlanta Public School System, in which the curriculum and instructional methods were analyzed with the view of giving the teachers a handbook. The studies, on which this handbook was based, were funded under the Elementary and Secondary Education Act of 1965 (ESEA), Title I, and subcontracted to The Center for Applied Research, Georgia State University. The handbook should be especially helpful for those persons who are not familiar with the principles and their applications.

This handbook is by no means all inclusive. Some of the basic ideas are presented; however, more specific information can be found in previous reports including "Teacher's Primer of Behavior Modification Techniques," Research and Development Report, Volume VI, Number 2, July, 1972 and "Achieving Academic and Social Objectives in Kindergarten Through Behavioral Analysis," Research and Development Report, Volume VI, Number 3, July, 1972.

Jarvis Barnes
Assistant Superintendent
for Research and Development

ABSTRACT

A Behavior Modification Program was implemented in four classes of a normal public school. These four classes encompassed 93 children ranging in age from 7 to 14 years. Included in the project was a second grade, a third grade, a sixth grade, and a seventh grade.

The primary focus of the program in the second, sixth, and seventh grades was upon increasing the academic performance of the children. In these three classes, a variety of methods for increasing academic performance was explored, including the use of student tutors and points for correct work. In the third grade the primary focus was upon control of extremely high rates of disorderly behavior through the use of a home-based management system.

The results revealed that in each of the three classes where academic performance was the primary variable, statistically significant differences were obtained on the *Metropolitan Achievement Tests* when compared to control groups. Further, all classes increased in number of academic "levels" passed. Additional experiments indicated that: (1) it was the point reinforcement system which controlled the children's rate of academic performance, (2) student tutors lead to increased levels of academic performance when compared to periods when no tutors were present, and (3) the academic performance of the tutors was facilitated by the process of teaching other children.

ACKNOWLEDGMENTS

The authors wish to express appreciation to the four teachers who participated in the 1971-72 Jessie Mae Jones School Behavior Modification Program: Ms. Cindy Kuhlman, (second grade); Ms. Brenda Grazier, (third grade); Ms. Gail Donner, (sixth grade); and especially to Ms. Elizabeth Tubbs, (seventh grade); who worked so closely with the Behavior Modification Staff throughout the year.

To our colleagues, Bill Skuban, Steven Garber, and George O'Neill, we extend appreciation for their assistance and participation during the many phases of this program. We enjoyed having them as co-workers and are grateful for their creative ideas.

Our sincerest thanks also go to Mr. Milton White, Principal of Jessie Mae Jones School, for his continued support and advice in our second year at his school. And lastly, to Ms. Catherine Houston, Secretary, who made our job much easier through her cooperation and assistance.

TABLE OF CONTENTS

TABLE OF CONTENTS (Contd.)

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

During the 1970-71 academic year, a Behavior Modification Program was conducted in one class at the Jessie Mae Jones School. The major objectives of that program were: (1) to establish control of the disorderly behavior in the class and (2) to increase the academic performance of the children. The results of that program indicated that, through a reinforcement program, the disorderly behavior of the class could effectively be brought under control and the academic performance of the children could be significantly increased. (The results of that program are described in detail in "Behavior Modification in a Seventh Grade Classroom," Research and Development Report, Volume V, Number 6, October, 1971, Atlanta Public Schools.) In spite of the success of this program, though, some significant problems remained unsolved. First, the new rates of academic performance generated by the reinforcement program resulted in more work than the teacher could effectively deal with; that is, the children were turning in more pages of work each day than the teacher could grade, an average of over 50 pages per day. Second, the question of how general the effectiveness of the procedures developed might be when applied to wide age ranges of children and different teachers remained unanswered.

The objectives of the program for the current academic year (1971-72) then, were to resolve the problems that the previous program had left unsolved and to expand the program to encompass more children. The classes chosen for this year's program included a wide age range of children (7 through 14 years old). This year's program involved attempts to develop grading procedures which would allow one teacher to grade all the work produced by the children and to do the grading in significantly reduced periods of time. Consistent with efforts to increase the amount of time available to the students for instruction, the use of student tutors was explored. Therefore, the 1971-72 program involved intensive efforts to develop procedures through which children could teach each other in certain subjects; thereby, freeing the teacher to spend more time in the subjects with which the children cannot help each other.

Other efforts in the 1971-72 program at Jessie Mae Jones School involved extending the subject matters studied from arithmetic and reading to other subjects and efforts to magnify the amount of work produced through the use of differential rates of reinforcement for varying amounts of work.

Although disruptive behavior was not the primary focus of this year's work, efforts were made to control disruption in all classes. This allowed the previous procedures (1970-71) to be tried with several different teachers.

II. DESCRIPTION OF THE PROGRAM

Personnel

A. Teachers. During teacher orientation week, four teachers -- one second grade teacher, one third grade teacher, one sixth grade teacher, and one seventh grade teacher -- volunteered to participate in the project. After consultation with the school principal, it was decided that all four teachers would take part in the project.

B. Aides. No teacher aides were used to implement or carry out any of the procedure.

C. Research Assistants. Two full-time research assistants were available for program design, consultation, and data analysis only.

Evaluation of the Children's Academic Status and Performance

In any behavior modification program which has academic performance as its basis, efforts must be made to monitor the academic output of the children as closely as possible. Without constant monitoring of the children's academic efforts, neither the teachers nor the staff would know the level of performance of the children. Daily monitoring provided constant evidence of the effectiveness of whatever set of procedures was in effect at the moment. Daily monitoring alone though is not enough, as efforts must also be made to monitor the overall effectiveness of the program; thus, providing evidence of the program's success or failure in generating increased academic

achievement. The two types of monitoring, daily and overall, provide data relating to: (1) the effect any given set of procedures was having upon the children's academic performance at any point in time and (2) the overall effect of the entire program upon the children's overall academic achievement between the start and completion of the program. The academic monitoring procedures used at Jessie Mae Jones provided such measures. The daily measures of number of pages completed correctly offered an index of the degree to which any given set of procedures motivated the children to engage in more or less academic work as compared to other procedures. The number of tests passed was a mid-point measure between daily performance and overall academic achievement, and the measures of number of levels passed and statistical analysis of Metropolitan Achievement Tests scores provided meaningful measures of the overall level of academic achievement for the class.

Obviously, though, in order for such measures as daily number of pages passed to be meaningful and relevant, the children must be working in academic material at their own level. To expect a child who is reading at the first grade level to complete pages in a seventh grade book would be sheer folly. Therefore, efforts first must be made to determine the academic level at which each child enters the program.

The following sections describe the procedures just discussed in detail.

A. Academic Placement of the Children. When the program was begun in each class, the children were tested in order to empirically determine the level at which they could successfully complete academic tasks. Placement tests were taken from the comprehensive tests in the back of the teacher's editions of the arithmetic series and were taken directly from the individual reading workbooks. All children were given the lowest test in each subject (first grade book in arithmetic and readiness level in reading) as their first test. All children who passed that test then were given the next higher test. This procedure was continued until all the children had reached a level at which they could not pass the test (score $\geq$ 70 per cent). Each child was then begun on the first page of the level he had failed. For example, if a child passed levels 1, 2, 3, 4, 5, and 6

of the reading series but failed level 7, he would be started on the first page of the level 7 workbook.

B. Daily Academic Performance. The daily academic performance was evaluated, in all cases, through the use of three different academic measures.

1. Pages: Each day a record was kept for each child of the number of pages he worked correctly in his subjects. This allowed a day-by-day sampling of the rate of performance of the children. The child was required to rework any page in which he did not score $\geq$ 70 per cent.

2. Tests: The children were required to take tests over the material just covered every time they completed a chapter. In reading, 15 pages were considered a chapter; and in arithmetic, the book was already conveniently divided into chapters. The number of tests passed under different conditions of reinforcement provided a valuable measure of motivation. If a child failed (scored less than 70 per cent) on a "chapter" test, the teacher discussed it with him, reviewed the areas in which he was weak, and then he was required to re-take the test.

3. Levels: The basic academic unit was the level. A level was defined by the publishers of the books utilized by the school. In the reading series utilized, the publishers divided the material into twelve books and defined each book as a reading level. In arithmetic, the publishers divided the material between first and seventh grade into seven books, each considered a level. The definitions offered by the publishers were retained, and each book was considered a level. These levels cannot be called "grade levels," as some grades in reading required the completion of two books or levels. Table 1 presents the books utilized in both reading and arithmetic. It should be stressed here that the academic materials utilized were those in use throughout the school; consequently, no special staff or materials were either needed or used.

TABLE 1

LIST OF TEXTBOOKS USED

Reading:	The Houghton Mifflin Readers. Levels 1 -- 12
Arithmetic:	Mathematics in Action. The American Book Company. Book 1 -- Book 7

Successful completion of a level, in either reading or arithmetic, required that the student turn in each page in the book at a criterion of 70 per cent correct, pass periodic "chapter" tests over the content of the book, and pass a final "section" test after completion of the book.

C. Overall Academic Achievement. The long term effects of the program were evaluated through the use of the Metropolitan Achievement Tests (MAT). Comparisons of the amount of gain achieved by the project classes from Fall to Spring on the MAT with the amount of gain for preselected control groups affords an index of the level of overall achievement of the children in the project classrooms.

Evaluation of Discipline Problems in the Classroom

Disruptive Behavior. Disruptive behavior, while not the major variable in any class but the third grade, was still of interest in all classes. This level of disruption was measured daily in all classes by observing rows of children. This daily measure of disruption provided an ongoing index of the level of disruption present in each class daily.

Teacher Behaviors

A. General Duties. The general duties of the teacher in the Jessie Mae Jones project classrooms involved all teaching, management, and program implementation. The research assistants did not help with

the actual implementation of any of these duties nor did they interact with the children in the classroom.

B. Academic Materials. The teacher made available all academic materials, in the form of books and workbooks, required for the children to perform their academic work.

C. Management Techniques. The basic management techniques for both encouraging academic performance and handling discipline problems involved the use of points. The children earned these points for behaviors judged correct or desirable and for correct academic performance. The points could then be exchanged for a wide variety of backup reinforcers.

Each child had a 3" x 5" card upon which appeared the child's name, the date, the number of points he currently had, and the backup reinforcers available. The children were given fresh point cards each day and were responsible for turning them in at the end of the day. If a child lost his card, he was given a new one; but the points that had been on his previous card were forfeited.

Periodically during the day, the teachers would put the points the child had earned on his point card, and in order to gain access to any of the reinforcers, the child had to present his card to the teacher. If the teacher saw that the child had enough points on his card to purchase the reinforcer, she simply punched the price of that reinforcer from his card. If he did not have enough points, he was sent back to his seat.

For reinforcing activities, the four teachers were divided into two "yoked" pairs; the second and third grade teachers were "yoked" together as were the sixth and seventh grade teachers. This "yoking" permitted the teachers to work together during reinforcing activities such as recess; thus, allowing one teacher to take charge of all the children from both classes who took part in the activity, and the second teacher to take charge of all the children from both classes who did not take part in the reinforcing activity. Such a procedure

made each "yoked" team independent and freed them from dependence upon aides or other outside personnel.

All classes that participated in the project had certain backup reinforcers in common, although the prices of those reinforcers varied from class to class.

"Game Room" was a large room equipped with many games, candy, and basketball goals. The children were required to purchase admission to the game room itself for a set number of points. Once inside, each activity or item that the child wanted cost extra points. The game room period lasted 40 minutes.

"Recess" consisted of a 30 minute period on the school playground. Admission to recess could be purchased for a set number of points. Once outside, items such as footballs, softballs, bats, and gloves were made available for varying prices.

"Trophies" were made available to all children whenever they passed a level. In order to purchase a trophy, a child had to pass the level and make his selection from a list of available trophies.

"Break" consisted of a five-minute period outside the room, during which the child could go to the bathroom or get water. Only one child at a time could "go on break," and if a child either overstayed his allotted five minutes or disturbed others, he was assessed a penalty of three times the price of the "break" itself.

Evaluation Procedures

A. General Description. Various strategies were available for use in comparative evaluations of applied programs and general research efforts. One approach was to compare the group in which some experimental program or special intervention was used with another group which was similar in all respects except that the experimental program (or techniques or exercises) was not used. Comparisons were usually made on the basis of some test or exercise that was common to both settings and yet flexible enough to allow distinction between the two groups. The test could be given before and after a

certain period of time, for example, a summer school program with a pre and post test evaluation. The test scores then could be evaluated in several ways, although the most common was through the use of standard statistical analysis. The inference was that any differences between the two scores most likely would be accounted for by the special program. There were, however, certain limitations to this method of comparison; for example, there was no way to identify which of the componenets of a special program accounted for the differences between the groups. One knows simply that there was or was not a difference but not the exact source or sources for any difference.

Another strategy for determining the effect of an intervention was to compare the group or the individual to itself before and after the intervention was used. Thus, any chnage in performance was most likely accounted for by the intervention. This strategy required no additional group to serve as a control group and was often referred to as a pre-post analysis.

In behavior modification programs, two other types of experimenta designs are often used. First, there is the ABA or reversal design. In this design, baseline measures are taken of some behavior -- disruption for example -- then the treatment program is implemented and measurement continues. Finally, the treatment program is withdrawn, but still the measurement continues. Any change which is not present in the first and last period but is present in the middle one then can be attributed to the treatment program. The other type of design used in behavior modification programs is the multiple baseline design. In this design, baseline measures are taken of several behaviors. Then the treatment procedures are successively applied to each behavior. The extent to which the level of performance of each behavior changes only when the treatment procedures are applied, is the extent to which it may be said that the procedures being used are effective. Following is a detailed description of the four types of evaluation procedures used in the Jessie Mae Jones project.

1. The ABA or Reversal Design. This design has been found to be effective in demonstrating experimental control in a wide variety of settings. Both Sidman (1960) and Baer, Wolf, and Risley (1968) have pointed out the advantages of such a design when compared to traditional statistical treatments. Utilization of this design involves the selection and definition of the behaviors to be studied. For example, number of pages of work in arithmetic done correctly each day might be the behavior of interest. First, baseline data would be gathered under certain specified conditions. This baseline would last for several days, in order to give an idea of the level of performance under whatever conditions are in effect at the moment. Next, some experimental manipulation would be carried out in an effort to determine the effects of the manipulation upon the behavior of interest. Finally, after several days of measurement of the behavior under the new conditions, the conditions would return to those in effect during the baseline.

As an example of this type of experimental design, if the number of correct pages in arithmetic per day under standard conditions was measured for 10 days and found to average 1.7 pages per day, this would constiute the baseline (or A phase). If the children then were given points for every page they worked correctly and found that for 10 days it averaged 3.8 pages correct per day, this would constitute the treatment (or B phase). Then, taking out the points and observing the behavior for another 10 days, it might be found that the average was only 2.1 pages correct per day. This would constitute the reversal (or A phase).

Under these conditions then, an effect upon pages done correctly per day has been demonstrated. If there is still some question as to whether or not the effect was due to the experimental condition, the points could be re-applied and withdrawn as many times as necessary in order to demonstrate adequately the effect.

2. The Multiple Baseline Design. This design was presented first by Baer, Wolf, and Risley (1968) and offers an excellent device for dealing with more than one behavior at once. In using this technique, a number of responses are defined and measured over time in an effort to provide baselines against which changes might be compared. Once the baselines are established, the experimental variable is then applied to one of the responses and changes in it are measured. If the experimental condition produced a change in the response, and little or no changes in the other responses, then rather than reverting to baseline conditions, the experimental procedure is applied to a second response. The experimental condition is then applied to a third response and so on. The objective is to demonstrate that the experimental variable is a reliable one, in that each response makes its maximal change only when the experimental manipulation is applied to it.

   For example, test scores in reading, science, and arithmetic might be measured over a period of 10 days. The experimental treatment, points for passing, then might be applied to reading. If reading increased, but not science or arithmetic, there is some indication that it was the experimental variable which caused the change. The next step would be to apply the experimental variable to a second subject. If that subject then increased, but the third, still untreated, subject remained down, then evidence for the strength of the experimental variable becomes even stronger.

3. Statistical Analysis. In this type of design, a comparative statement is made based upon the statistical probability of certain changes occurring. For example, it would be possible to compare the performance of two groups on a test, one of which had an experimental program in its class, the other taught with "standard classroom techniques." This type of

comparison is referred to as between groups. The major assumption of this comparison is that any differences between the experimental and control group will be due to the presence of the experimental program, the independent variable. A second type of statistical probability statement is typically called the within group comparison. For example, if a class was given a test in August and some independent variable brought into effect until June, when a second test was given, it would be possible to compare the two scores to each other and make a statement as to the probability of such a change occurring by chance. In statistical designs the between group comparison is preferable to the within group comparison as the within group comparison does not exclude time alone as an independent variable.

4. Pre-Post Analysis. This design is often similar to the statistical analysis in that the experimenter compares data at two discrete points in time. He may or may not utilize a statistical treatment although this is the usual procedure. Another use of this technique would involve determining, for example, at what level a child was reading in August and determining again in May the level at which the child was reading. Comparing the reading levels at the two points in time would then give a measure of number of levels increased from one point in time to the next.

SECOND GRADE

Introduction

The final procedures utilized in the second grade were the end results of a series of problem-oriented experiments. The major problem remaining from the previous year's program (1970-71) involved grading the children's work. It was necessary that this problem be resolved before progress beyond that previously attained could be achieved.

In order for the teacher to grade completely all the work produced by the class (large amounts of time were required) thus, consuming periods of time which the teacher might devote to other activities such as instruction. The solution to the problem is complex but would seem to center around one of four alternatives: (1) having the children produce less work, (2) providing the teacher with an aide or other personnel for assistance with grading, (3) requiring the teacher to spend large amounts of after school time for grading, or (4) not grading all the work produced by the children. The first solution, having the children produce less work, is clearly unacceptable for a number of reasons although the type of work produced might be re-arranged so that more oral work, which does not require later grading, is produced. Even oral work has its problems though, as the very nature of such work requires a one-to-one ratio between teacher and pupil; thus, consuming large amounts of time. The second solution, providing aides, is unacceptable from a cost point of view. No public school can afford to hire aides to do nothing but grade children's work. The use of student graders may prove an acceptable alternative, but even here students cannot be expected to spend the long amounts of time required for grading all the work produced. The third solution, requiring the teacher to grade after school is unacceptable for two reasons: (1) the teachers will be very resistant to such a procedure and (2) immediate feedback is much more effective than the delayed feedback which would be created by such a situation. The fourth alternative, not grading all the work produced by the children, appears at first glance to be no more promising than the other three. If a child turns in five pages and the teacher grades only one or two, then he gets no feedback on his performance on the other pages, and it is very possible that he has made mistakes on those ungraded pages which need to be corrected. Another use of this alternative, though might involve grading only part of each page, as pages usually contain material which is related. The problem of time necessary for grading required resolution early in the program. The first experiment presents an effort to resolve this problem.

Experiment I -- Grading Procedures

Children

The entire second grade was included in the Behavior Modification Program. There were 19 black children whose ages ranged from 7 to 9 years with a mean of 7.6 years.

Response Definition and Recording

The response here was performance on reading workbook pages and skill sheets which accompany the workbook. A per cent correct for each page attempted was computed.

Experimental Design

The experimental design for this study employed statistical measure to determine the extent to which the scores obtained by the new method would correlate with those obtained by the old method.

Further, a pre and post series of observations were employed in an attempt to determine the effects of the new grading procedures upon the performance of the children.

Procedure

The pages turned in were cast into one of two groups: (1) any page on which the child had gotten at least 70 per cent of the problems correct was considered passed and (2) any page on which the child scored less than 70 per cent was considered failed. When the new procedures were started, the teacher (or grader) was instructed to select one problem from each page and grade that problem. The teacher (or grader) was instructed to select the problem for grading which she felt was most representative of that entire page. If the child passed that problem, he was considered to have passed the page; if he failed that problem, he was considered to have failed the page. In order to check the reliability of "quick" grading, the Behavior Research Staff graded the first 750 pages turned in under this method in their entirety. This allowed a comparison between the number of children

who passed using the "quick" grading technique and the number who would have passed using standard grading procedures.

Also, a record was kept of the average per cent of pages passed for the class when both "quick" and "complete" grading techniques were used. This percentage was taken from the complete grading done after school by the Behavior Research Staff. This procedure afforded an opportunity to examine the effects of "quick" grading upon the children's performance.

Results

The results of a page by page comparison of 750 pages graded under both "complete" and "quick" grading techniques revealed that in 94 per cent of the cases there was agreement between the "quick" and "complete" grading techniques as to which children would have passed or failed a given page.

Further, when the class was operating under the "complete" grading technique, the class passed an average of 88 per cent of the pages turned in. When the "quick" grading technique was employed, the class passed an average of 86 per cent of the pages turned in. These percentages are based on the "complete" grading done by the Behavior Research Staff, and indicate that the performance of the children did not suffer from implementation of the "quick" grading procedures.

Discussion

These results indicate that it is possible to develop grading procedures which require less of the teacher's time than grading each problem submitted by the class. They are also designed so as to result in grades that are truly representative of each child's performance. The importance of developing grading techniques such as these can only be made clear by re-stating the fact that the advancement of the entire program had to await the development of the new grading procedures.

In spite of the savings achieved by procedures such as have been described, the teacher may still find herself overwhelmed by the enormity of her burden. The task of instructing children of widely varying levels of ability is a complex problem, both logistically and methodologically. This problem is

typically handled by dividing the children into groups according to their levels of ability. When the class is divided into groups, the teacher can base her level of instruction around the academic level at which that group is functioning. Although this solution to the problem of instructing children of widely ranging abilities has much to recommend it; the teacher can only work with one group at a time, thus, leaving the other groups to work on their own with little or no teacher assistance. It seems likely though, that there are certain academic exercises that children can teach each other, for example, vocabulary recognition. In order to teach a child a word, it is not necessary that one be a teacher; it is only necessary that one know the word to be taught. Therefore, in an effort to solve the problem of insufficient time for the teacher to instruct all groups in all skills, a system of tutors was devised. The next experiment describes the implementation and results of a system for the use of older children (sixth and seventh graders) as teachers for younger (second graders) children.

Experiment II -- The Effects of Sixth and Seventh Graders as Tutors in a Second Grade Classroom

Children

The subjects were the 19 second grade children already described plus sixth and seventh graders, members of two of the other experimental classes, who served as the tutors.

Response Definition and Recording

Three dimensions of reading were selected as the target behaviors. They were:

A. Vocabulary Recognition. The verbalizing of basic sight words when presented.

B. Oral Reading. Reading out loud in stories in the basal reader.

C. Written Work. Responding in writing to written questions found in the workbook for the reading series.

The last of these dimensions leaves an enduring record of the child's performance. Recording was accomplished by the research assistant going to the workbooks each day and computing a per cent correct for each child (i.e., number of pages done correctly/total number of pages attempted.) The other dimensions required direct recording which was accomplished by having the teacher take samples of each of these behaviors during her reading groups. This experimental requirement caused little problem due to the fact that the teacher was already in the habit of introducing new vocabulary words and having the children read orally during their groups in the morning. For vocabulary she presented each child with five words written on index cards. These words included the three new words which were introduced that day and two "old" words from the previous day's list. The teacher recorded on a data pad correct and incorrect responses and a per cent correct was calculated from this.

For oral reading, the teacher had the students read in turn, pages from the basal reader, and using a wrist counter she recorded the number of mistakes per page. She then recorded this data on paper. By going back and counting the number of words in each page, the Behavior Research Staff was able to calculate a per cent correct per page per child.

Experimental Design

The design used here was a combination of ABA (reversal) and multiple baseline approaches. An ABA was run on each of the three dimensions. Simultaneous recording of the other two response classes was continued during these reversals. The design is shown in Table 2.

The independent variable under manipulation was the added teaching by the sixth-seventh grader of the particular dimension under study. For example, in the case of vocabulary a student teacher was present in block 1, absent in block 2, and present again in blocks 3 through 7. In the B phases where there was no student teacher present, the group that would have been working with that teacher was merely given the stimulus materials and told to work on them as hard as they could.

The purpose of this design not only was to determine the effects of student teachers, but also to determine the interaction effects among the various dimensions of reading.

TABLE 2

EXPERIMENTAL DESIGN

Blocks of Time	1	2	3	4	5	6	7
Number of Days in Each Block	6	10	5	7	6	6	6
Vocabulary	A	B*	A	---------->			
Oral Reading	<--------		A	B*	A	------->	
Workbook	<----------				A	B*	A

* No tutors in any of the B blocks, tutors in all other blocks.

Selection and Definition of Backup Reinforcers

An effort was made to select backup reinforcers which were natural events in the school environment, such as recess, or were relatively inexpensive such as small pieces of candy. Many items such as recess and game room could be purchased only at set times each day, but other reinforcers, such as break, were available at any time during the day that the child chose to avail himself of the reinforcer. In this class, points were re-cycled every day, that is, the children could not save points from day to day. As game room always took place during the last forty minutes of the day, ample opportunity was provided for the children to exchange all their points during that period. Table 3 presents a summary of the items available for exchange and the cost of each item.

TABLE 3

LIST OF BACKUP REINFORCERS AND THEIR COST

Item	Price
Trophy	100 + Pass level
Game Room	40
Recess	20
Restroom	1
Candy	3
Cookies	4
Comics	3
Magazines	4
Cap Gun (rent 10 min.)	5
Caps (for gun)	5
Basketball (rent 30 min.)	3
Baseball bat (rent 30 min.)	3
Ball (rent 30 min.)	3
Hats	2
Rings	15
Creepy Crawlers	10
Hair Ribbon	5
Dolls (rent 30 min.)	5
Bolo Bat	5
Toy Cars (rent 10 min.)	4
Skates (rent 30 min.)	10

The points were exchanged in the manner described in the "General Procedures" section.

Procedure

At the beginning of each day, the sixth and seventh grade teachers selected at random students to go to the second grade class to tutor. This choice was not based on academic excellence but rather an attempt was made to allow every child to be a tutor. The only criterion used in selection was that no student be allowed to tutor more than once a week. This was necessary because tutoring required most of the morning and some of the school officials had expressed concern that the time these children spent away from their classes might be harmful to their studies. As it turned out, the total number of students in the two upper classes was over forty and this meant that each child missed class only once every ten days.

Tutors reported to the second grade class at 9:00 a.m. They were introduced to the younger children as the "assistant teachers" and then taken by the teacher to her desk for ten minutes of briefing. One sixth-seventh grader was assigned to each of the following duties.

A. Teaching Vocabulary. The tutor was given a stack of flash cards containing the new words for the day plus any previous words given that week. These words were taken from the suggested list in the teacher's guide of the Houghton Mifflin series and three new words were added to the cards each day (this was the prescribed rate for introduction of new words). The teacher had already put these words on the board and had introduced these words in a traditional fashion prior to the tutors entering the room. The tutor was instructed to meet with his students and to flash the cards quickly, giving each child a chance to identify the word. Moving clockwise, he was to continue until one child responded correctly. He then was to give that child one point for his correct answer. After the correct response was made, all of the children repeated the answer. Then the tutor proceeded to the next word.

B. Teaching Oral Reading. The tutor was given the basal reader for the series and told which pages were to be covered for the day's lesson. He was instructed to have his students read, in turn, successive pages from the assignment. If the student had trouble with a word, the tutor was told to help him figure out the word. One point was given for each page read with less than four mistakes.

C. Teaching Written Work. The second graders had been placed in the proper level of the reading series workbooks. They received one point for each page done correctly (70 per cent). The tutor was told to help the children with their workbooks by explaining instructions and teaching principles but not by giving answers.

D. Grading. One tutor was given the teacher's answer books for the workbooks and told to grade the books as the children finished. The tutor not only marked the problems in the workbook, but also recorded the results on separate data sheets. These sheets enabled the

teacher and the research staff to determine which pages the child had attempted and which he had passed. Also, the tutor gave the children the appropriate number of points for their work.

The class was divided into four groups of five students each. The desks of each group were arranged to form a "U" (see Table 4). A chair was placed in the mouth of each "U." The regular teacher sat with one group while a tutor sat with each of the other three groups. The fourth tutor sat at the teacher's desk and graded workbooks. The teacher set a standard kitchen timer for twenty minutes. At the end of each time period, the regular teacher and the tutors moved counterclockwise one group. There were four such periods. Therefore, by the end of the reading session, the teacher and each tutor had met with each of the four groups. Thus, the teacher was able to cover all of the areas prescribed by the teacher's guide. At the same time, she was able to monitor the efforts of the tutors who were giving each student extra practice in each of the dimensions specified.

TABLE 4

DIAGRAM OF GROUP TEACHING IN THE SECOND GRADE

N = 21

Period 1

Group 1	Tutor A	Group 3	Tutor C
Group 2	Tutor B	Group 4	Teacher

TABLE 4 (Cont'd.)

Period 2

Group 1	Tutor C	Group 3	Teacher
Group 2	Tutor A	Group 4	Tutor B

Period 3

Group 1	Teacher	Group 3	Tutor B
Group 2	Tutor C	Group 4	Tutor A

Period 4

Group 1	Tutor B	Group 3	Tutor A
Group 2	Teacher	Group 4	Tutor C

Tutor A - Vocabulary

Tutor B - Workbook

Tutor C - Reading

Although level of disruption was not measured during this experiment, the teacher had a "quiet point" system based on the one designed for the 1970-71 program at Jessie Mae Jones. This quiet point system allowed the teacher to maintain control of disruption with minimal difficulty.

Results

The findings of this study are shown in Table 5. The nonsquared numbers indicate, for each dimension, those time blocks in which the treatment was in effect; and the numbers enclosed by squares indicate reversals, when no tutors were present. Looking first at vocabulary, it was evident that the removal of the tutor had a disasterous effect on the performance of the second graders. The per cent correct on vocabulary tests decreased from 78.7 per cent to 49.9 per cent. This difference was significant at the .01 level. A return of the treatment brought the performance back up to 81.2 per cent and this was also significant at the .01 level. Following block three there was a downward trend in vocabulary performance but all of these time blocks were still significantly above the reversal phase.

TABLE 5

ANALYSIS OF THE EFFECTS OF TUTORS TEACHING ON THREE DIMENSIONS OF READING

(Performance in Per Cent Correct)

Blocks of Time	1	2	3	4	5	6	7
Number of Days in Each Block	6	10	5	7	6	6	6
Vocabulary	78.7**	49.9	81.2**	74.2	70.3	71.2	71.5
Oral Rading	73.0**	44.7	71.8**	62.6	73.0*	71.5	69.7
Workbooks	77.2	70.7	73.6	89.1	88.0	72.0	95.0**

* $p > .05$.
** $p > .01$.

[ ] No tutor.

Discussion

The results of this experiment demonstrate conclusively that sixth and seventh graders can effectively serve as tutors for second grade children. However, it is not always possible to get older children to serve as tutors. School activities which involve only the older children are often planned, thus, interfering with the second grade teacher's program for that day. Further, teachers may be reluctant to allow their children to leave their rooms to serve as tutors in other classes on any regular basis. An effective solution to this problem would be to utilize the children in each class as tutors for each other. The next experiment describes the use of second graders as tutors for other second graders.

Experiment III -- Second Graders as Tutors for Other Second Graders

Children

The children were the same 19 second grade children described in the previous experiments.

Response Definition and Recording

The responses utilized and the method of measuring them were identical with those used in Experiment II.

Experimental Design

The experiment design for this section was a simple pre-post design with the grade (sixth and seventh or second) of the tutors as the independent variable.

Procedure

The procedures employed in this experiment correspond exactly with those utilized in Experiment II, except for the use of second graders as the tutors rather than sixth or seventh graders.

Results

The performance of the second graders versus sixth-seventh graders in the three dimensions of reading is shown in Table 6. It was obvious from these results that functionally there was little difference between the effectiveness of the two types of teachers. In vocabulary, second grade teachers actually helped their classmates to perform one per cent better. In reading, sixth-seventh grade teaching resulted in four per cent better performance. Finally, in workbooks the second grade helpers did 5.4 per cent better than their sixth-seventh grade counterparts. Overall then, second grade teachers did 1.2 per cent better than the sixth-seventh grade teachers. None of these differences were significant.

TABLE 6

RELATIVE EFFECTIVENESS OF CHILDREN AS TEACHERS OF SECOND GRADERS

N = 19

2nd Grade	7th Graders as Teachers (Per Cent)	2nd Graders as Teachers (Per Cent)
Vocabulary	72.2	73.2
Reading	76.2	72.2
Workbooks	81.8	87.2
TOTAL	76.7	77.5

Discussion

The results of this experiment demonstrated empirically that in certain subjects, second graders can effectively serve as tutors for other second graders. The fact that children can teach other children is, however, not a valid reason alone for using tutors. If it cannot be demonstrated that the tutors themselves derive some benefits from the process of helping other children; then the teacher ethically cannot deprive any child of time, when he might be learning himself, just to help her teach another child. The next experiment was an attempt to resolve this problem.

Experiment IV -- Some Effects of Tutoring Upon Both Tutors and Their Students

Children

The children were three second graders; two males, Anthony and Kenny and one female, Sylvia. Also, two other male second graders, Stanley and John, were selected to tutor the first two boys. The teacher worked with Sylvia. The three children, Anthony, Kenny, and Sylvia, were selected because their performance during the previous experiments had fallen well below that of the remainder of the class. These three children averaged only 39 per cent of their vocabulary words correct, even when being tutored, and only 5 per cent when not being tutored. The remainder of the class had averaged 85 per cent and 49.9 per cent during the same two periods. It was believed; therefore, that these three children would benefit from individual tutoring. The two children who were to serve as tutors, John and Stanley, were selected because their performance approximated the class mean.

All of the subjects appeared to be healthy and normal except Sylvia who had been near mute all year. The teacher had not been able to get her to speak more than a few words and none of the staff recalled her ever interacting or speaking to her peers. She spent great amounts of time staring off into space and had done little or no written work all year. The teacher was concerned that Sylvia was retarded and, consequently, had referred her for testing.

Response Definition and Recording

The response here was performance in vocabulary. The data was recorded just as in Experiments I and II (i.e., performance on the five word vocabulary test given by the teacher in the reading groups). All children still were given the same test and the performance of the three target children and the two tutors merely were calculated separately for purposes of this experiment.

Experimental Design

In this study a multiple baseline across subjects was used. The independent variable, the presence of the tutors, was applied successively to each subject.

Procedure

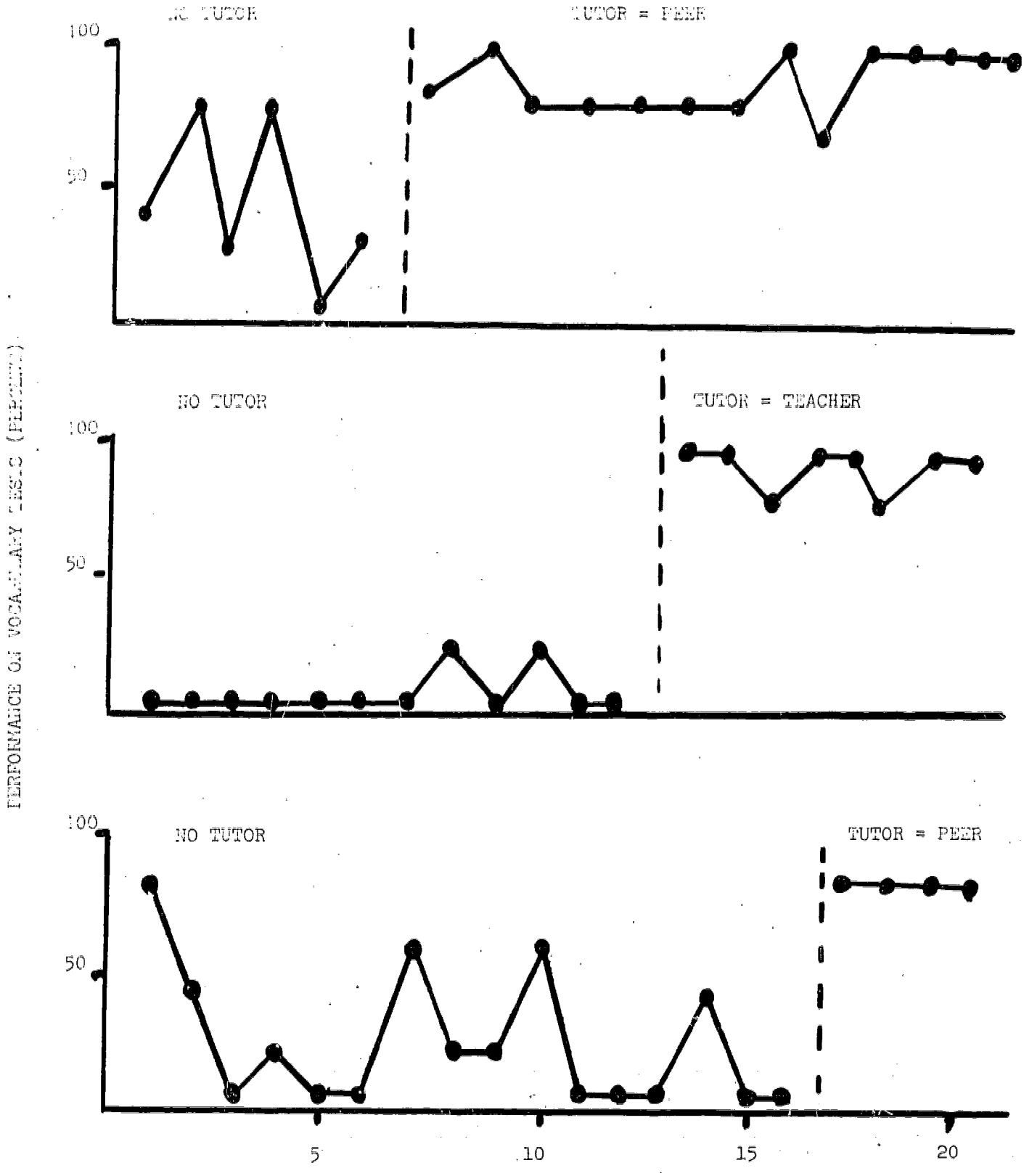
Each morning before reading groups started, the teacher called the children to be tutored and the tutors to her desk. She gave the tutors a separate set of flash cards containing the words for that day. The tutors were instructed to take their students over to the corner of the room and drill them on their words. If the tutor did not know a word himself, he was told to ask one of the children in the most advanced group of the class. The tutors also were told that if at the end of twenty minutes, their student had learned the words sufficiently (i.e., not more than one mistake), then they and their student would each receive a bonus of five points. At the end of the twenty minutes, the teacher sat with each pair and listened to presentation of the words. If in that recitation the student met the criteria, they both received praise and were given the bonus. If, however, he did not perform well enough, the teacher told both children she was sorry and reminded them that they had another chance later in the groups to earn their usual points for vocabulary. The teacher herself worked with Sylvia individually using the same procedure.

Results

Figure 1 shows the performance of each student before and after tutoring began. Anthony's scores increased from an average of 40 per cent correct during baseline to an average of 88.6 per cent after he was tutored by Stanley. In systematic replication, Kenny's performance increased from an average of 21.2 per cent to 100 per cent after he was taught by John. Sylvia also made dramatic improvement under the tutelage of the teacher. Her scores increased from an average of 3 per cent to 95 per cent. This is especially impressive when it is recalled that up to this point Sylvia had bearly spoken.

FIGURE 1

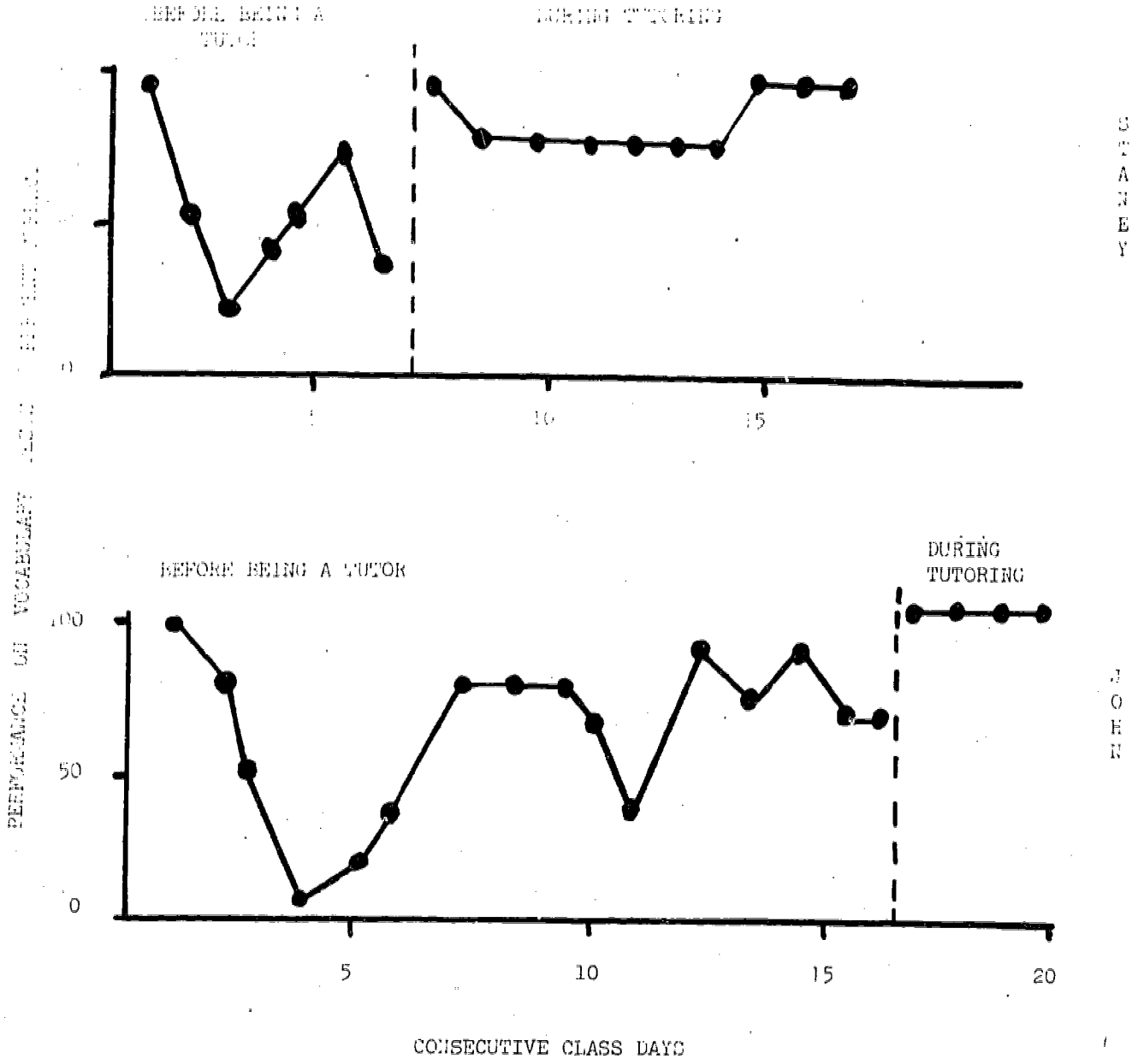
THE EFFECTS OF INDIVIDUAL TUTORS ON CHILDREN



In Figure 2, the effects of the tutorial experience on the tutor rather than his student is shown. Stanley was performing at the 56.7 per cent level before he began teaching Anthony. After his new role began, he scored 85.5 per cent on his own vocabulary test. Similarly, John was averaging 63.7 per cent prior to being made a tutor and 100 per cent afterwards.

FIGURE 2

THE EFFECTS OF BEING A TUTOR ON THE TUTOR HIMSELF



Discussion

The experiments presented so far have demonstrated techniques for (1) providing the teacher with assistance in instruction through the use of student tutors and (2) reducing the amount of time required for grading the students' work with no loss in efficiency of grading. One way to assess the overall impact of such procedures would be to measure the amount of time the teacher spent each day engaged in various activities, both before and after the implementation of such a program. The next experiment describes such an attempt.

Experiment V -- The Allocation of the Teacher's Time

Subject

The subject of this experiment was the regular teacher of the second grade. She was a 23 year old Caucasian female. The 1971-72 school year was her second year of teaching.

Response Definitions and Recording

A. Teaching Time. This was the amount of time that the teacher actually spent engaged in instructional activities with either one child or a group of two or more children.

B. Grading Time. This was the amount of time that the teacher spent in grading or marking the academic work produced by the children.

C. Management Time. This was the amount of time that the teacher spent each day in managing the class. It included such activities as disciplining the children and maintaining order, seeing that all the children had pencils, and getting them lined up to go outside. After implementation of the Behavior Modification Program, it included such activities as marking the point cards.

An observer recorded the amount of time the teacher engaged in these various classes of behavior for two hours per day, using a

stopwatch to measure the time. Periodically, a second independent observer would take measurements at the same time in an effort to obtain a reliability check on the measures. The reliability of the two sets of independent observations never dropped below 92 per cent and the mean was 94.6 per cent.

Experimental Design

The design utilized in this experiment was a simple pre-post design in which certain of the teacher's behaviors were measured before the Behavior Modification Program was implemented and the same behaviors were measured after the program was implemented.

Procedure

The procedure for this experiment simply involved the teacher engaging in her daily tasks before and after implementation of the Behavior Modification Program; and the measurement, by the observer, in which of the three classes her behavior fell.

Results

The results of this experiment are depicted in Table 7. It is revealed that the teacher's teaching time increased from around 10 per cent to more than 70 per cent following implementation of the Behavior Modification Program and tutorial system; thus, providing her with much more time for instructing the children. As the tutors were doing the grading after the program began, her grading time fell to zero; and, in spite of the management procedures required by the point system, such as marking point cards, her management time also declined.

TABLE 7

ALLOCATION OF THE TEACHER'S TIME UNDER TWO DIFFERENT SYSTEMS

(N = 19)

Tasks	Per Cent	
	Before Tutorial System	During Tutorial System
Teaching Time	10	76
Management Time	50	26
Grading Time	40	0

Discussion

This experiment has demonstrated that the grading and tutoring procedures described in the previous experiments have greatly increased the amount of time available to the teacher for actually instructing the children and has reduced sharply the amount of time she had to spend in grading and management activities.

Thus, experiments have been presented which have described a program through which: (1) the day-to-day academic performance of the children has been facilitated, (2) the amount of time required for grading has been reduced, and (3) the amount of time available to the teacher for actually teaching subject matter has been increased. An, as yet, unanswered question involves the effectiveness of such a program upon the overall academic achievement of the children. The following study describes one method for measuring such long-term effects.

Experiment VI -- Levels Passed

Children

The subjects for this experiment included the entire second grade class of 19 children that has been described previously, plus one child who withdrew in February for a total of 20 children.

Response Definition and Recording

The definition of the level has been previously given in the "General Procedures" section and will not be repeated here. The recording procedure consisted of simply keeping the check sheets on which the teacher recorded each child's successful completion of pages and levels. As reading was the primary subject matter within the second grade, materials were made available that allowed the children to pass levels only in reading, but not in arithmetic. The class worked on arithmetic daily, but the primary focus was upon reading.

Experimental Design

The experimental design consisted of a simple pre-post measure of the number of levels advanced by the class of 20 children during the 100 academic days that the project was operating in the second grade class.

Procedure

As previously described, the children had one 20-minute period each day during which they worked in their workbooks under supervision. They were free also to work in their reading workbooks during any spare time they might have. For each page that they completed correctly ($\geq$ 70 per cent) earned the student 100 points and the right to purchase a trophy of his choice.

After successfully completing a level, the student was then provided with all the materials necessary for his work on the next level, and he was allowed to begin work.

Results

The results of this experiment indicated that during the 100 academic days the project was operating in the second grade, 36 reading levels were successfully completed. Table 8 shows that only one child (Sylvia) failed to complete at least one level successfully, while 17 children completed two or more levels. The overall rate of passing levels for this class was one level passed every 2.78 days.

TABLE 8

NUMBER OF CHILDREN PASSING LEVELS*

Number of Levels Passed	No. of Children
Number of children passing one level	2
Number of children passing two levels	14
Number of children passing three levels	3
Total Children	19
Total Number of Levels	36

* 100 academic days.

Discussion

The results of this experiment are very clear and indicate that the children passed many levels in a relatively short period of time. However, while the children increased in number of levels passed, as the procedures required that they work each page in their workbooks before being allowed to take level tests, it might be questioned whether or not these results are an inevitable consequence of such procedures and nothing else. That is, the results in terms of number of levels passed might simply reflect the emphasis of the program and not be reflected in any other measures. Consequently another more standardized measure would provide a meaningful index of both the overall effectiveness of the program and the validity of levels as an index of overall performance. The last study provides such a measure.

Experiment VII -- Statistical Analysis of the Metropolitan Achievement Tests

Children

The subjects for this experiment involved two classes, the experimental group and the control group. The experimental class was the second grade class consisting of the 19 children previously described. The control class was pre-selected from another neighborhood school and involved 17 children.

Experimental Design

The experimental design utilized for this experiment involved a comparison between the amount of gain manifested by the experimental group of the Fall and Spring administration of the Metropolitan Achievement Tests (MAT), and the amount of gain manifested by the control group during the same period of time on the same test.

The independent variable was the presence of the reinforcement system in the experimental class and the dependent variable was the scores of the class on the MAT.

Procedure

The procedure involved the administration of the MAT at the two intervals (Fall and Spring) and the implementation of the reinforcement system already described in the experimental class during the interim. The control class was taught through the use of standard classroom techniques.

Results

The results of the analysis of the MAT yielded significant increases in both reading (t [32 df] = 2.499, $p > .04$ one tailed) and arithmetic (t [28 df] = 3.893, $p > .005$ one tailed) when the experimental group was compared to the control group. These results supported the hypothesis that the reinforcement system leads to increased academic performance.

General Discussion

The results of the experiments conducted in the second grade indicated that through the use of tutors' dramatic saving of teacher time can be made, and that concurrently the children's academic performance can be facilitated to a significant degree. The amount of time that the teacher had available for actually teaching the children was a critical variable in any class, as it was impossible for a teacher to instruct when she had to be managing the class or grading their work. This nonteaching time might be compared to what computer experts refer to as "down time." Any program which can increase the amount of time a teacher has available is extremely valuable and demands

strong consideration for widespread implementation. The grading and management procedures described for this class have experimentally demonstrated their worth in that they can increase the amount of time a teacher has available for instruction.

The use of tutors in this class demonstrated that the student-instructor ratio can be dramatically reduced through the use of available resources. Children can instruct each other, in certain subjects, and both tutor and tutee benefit from this relationship. The fact that the tutors benefit from the teaching process is critical, as a teacher could not afford to have tutors just to relieve her own burdens. If the tutors themselves were not learning from the process of being teachers, then it would be difficult to justify depriving them of the opportunity for instruction. The fact though that the tutors not only greatly facilitate the performance of the students they instruct; but have their own performance enhanced, lends strong weight to the argument that tutors should be used.

It must be pointed out that in all the experiments in which tutors were used, the children were receiving reinforcement in all conditions. That is, the children were earning points for correct academic performance even when they were not being tutored and, therefore, were motivated to perform as well as they possibly could. Without the tutors, however, the children simply did not have the information and skills necessary for high performance. If a person does not know how to add, no amount of potential reinforcement is going to cause him to suddenly be able to add. The reinforcement may, however, encourage him to try and learn how to add; but if no teachers or materials are available, the person still will not be able to add. In this class, the tutors provided the avenue through which the motivation created by the points could actually benefit the children.

The fact that the experimental class had significantly greater gain on the Metropolitan Achievement Tests (MAT) lends credance to the assertion that the system in effect in this second grade is significantly more effective than are standard classroom procedures. Further, the significance on the MAT lends strength to the concept of levels as advanced in this project. The fact that the children in this class increased 36 levels in

100 academic days is very impressive; but when combined with the MAT scores, the validity of this measurement is enhanced.

In summary, then, the results in this class have demonstrated effective techniques for increasing the amount of actual teaching time a teacher has available can be dramatically increased. Further, the use of student teachers can maximize the academic performance of both the students and their tutors.

THIRD GRADE

Introduction

In the third grade, as in the second grade, the procedures developed were the result of efforts to solve the problems present in this class. Unlike the second grade, though, the major problem in the third grade was disruptive behavior.

The program was not implemented in the third grade class until the middle of February, 1972. Initially, the teacher felt that she did not have any disruption problems and she wanted to teach academics in a manner that would have made implementation of a reinforcement program somewhat difficult. Consistent with the view that only the teacher can best decide what procedures will be most beneficial to her, the Behavior Research Staff assisted her on a consultation basis only until February. At that point, she requested immediate help with disruption as she felt her class had become unmanageable. Initial observations on the part of the Behavior Research Staff validated the teacher's statements. What had been a very orderly class during the first two months of the academic year had now become the most disruptive class ever witnessed by the Behavior Research Staff. Accordingly, the first efforts focused upon elimination of the disruption. This choice was made because it was quite obvious that no learning could take place in the class as it was presently constituted; disruption had to be reduced before academic performance could be accelerated. Initially, the procedures developed during the 1970-71 year at Jessie Mae Jones School were employed. Surprisingly, however, these procedures were only partially effective. Reinforcing good behavior reduced the level of disruption initially but failed to keep it at a constant low level.

Consequently, after consultation with the principal and the school counselor, a home-based program was designed. All the parents of the third grade children were sent a letter requesting that they attend a special parents' meeting at J. M. Jones School. The importance of the meeting was stressed by notifying the parents in the letter that their children would not be allowed to return to school unless they attended the meeting. The following study describes the procedures developed and their results.

Experiment I -- Elimination of Disruption Through the Use of a Home-Based Management System

Subjects

The subjects in this class were 23 black children who ranged from 8 to 10 years old, with a mean of 8.6 years.

Response Definitions and Recording

Disruptive behavior was defined as follows:

A. Noise behaviors: loud talking, singing, clapping, shouting out the teacher's name, and loud laughing.

B. Gross motor behaviors: out of seat without permission, fighting, slapping, foot stomping, and throwing objects. All behaviors other than those listed were scored as nondisruptive.

The measures of disruptive behavior were taken each class day, usually between 10:00 a.m. and 11:30 a.m. The children were observed by rows, and each row was treated as one unit. That is, the class was seated in four rows, so the observer could watch the entire row. If one child was disruptive during the time his row was observed, the entire row was marked as disruptive for that interval. Such a procedure does away with the necessity for choosing target children for observation and allows the observation of an entire class.

The observation procedure was as follows: the observer would watch row 1 for two minutes, then shift to row 2 for two minutes, then row 3, then row 4. The observer then shifted back to row 1 and restarted the cycle. The two

minute observation periods were divided up into 10-second blocks so that there were 12 intervals during any two minute period when a row would be marked as either disruptive or nondisruptive. Periodic reliability checks taken by a second observer revealed that when reliability was checked interval by interval, the two observers' percentage of agreement was never below 93 per cent and had a mean of 96 per cent.

Experimental Design

The experimental design utilized in this experiment was an ABA or reversal design.

Selection and Definition of Reinforcers

The reinforcers utilized in the third grade were identical with those utilized by the second grade (see Table 3).

Procedures

At the special parents' meeting, the principal, the counselor, and the Behavior Research Staff explained the program to the parents and told them of the problems. The help of the parents was then requested. They were told that they would, starting on the following Monday, receive a daily "good behavior" letter if their child was good in school that day. It was stressed that they should ask their child for his letter daily and they were told that if he had no letter it would only be because he had been a severe problem in school that day. They were told that the letters would be most effective if they showed their child how proud they were of him when he had one, and if they withheld some privileges such as playing outside when he did not have a letter. In practice the teacher also called the parents of the children who did not receive "good letters" at night to remind them that their child had not earned his letter that day.

Further, the parents were asked for their permission to keep their children up to thirty minutes after school whenever they were disorderly. All the parents were very enthusiastic about the letters, and they all readily agreed to allow their children to remain after school.

On Monday the following program was implemented: the children could earn quiet points for each 15-minute period they were quiet. If a child was called down once during a 15-minute period he lost his point for that period. A child who was called down twice during a single 15-minute period lost all the quiet points he had earned up to that point, and a child who was called down three times lost his "good letter" and had to remain after school for thirty minutes.

The periods were timed through the use of a kitchen timer and the teacher had a sheet and clipboard. Each sheet had every child's name on it and was divided into columns, each column corresponding with a 20-minute time period. Whenever a child was disorderly, the teacher simply placed a check by his name in the column corresponding with that time. This allowed a permanent record of disruption.

During the second phase only the letters and after school time were altered. During this period the children received their "good letters" when they came in in the morning and were told that there would be no after school time that day.

During the final phase, all conditions were reversed to those in effect during the first phase.

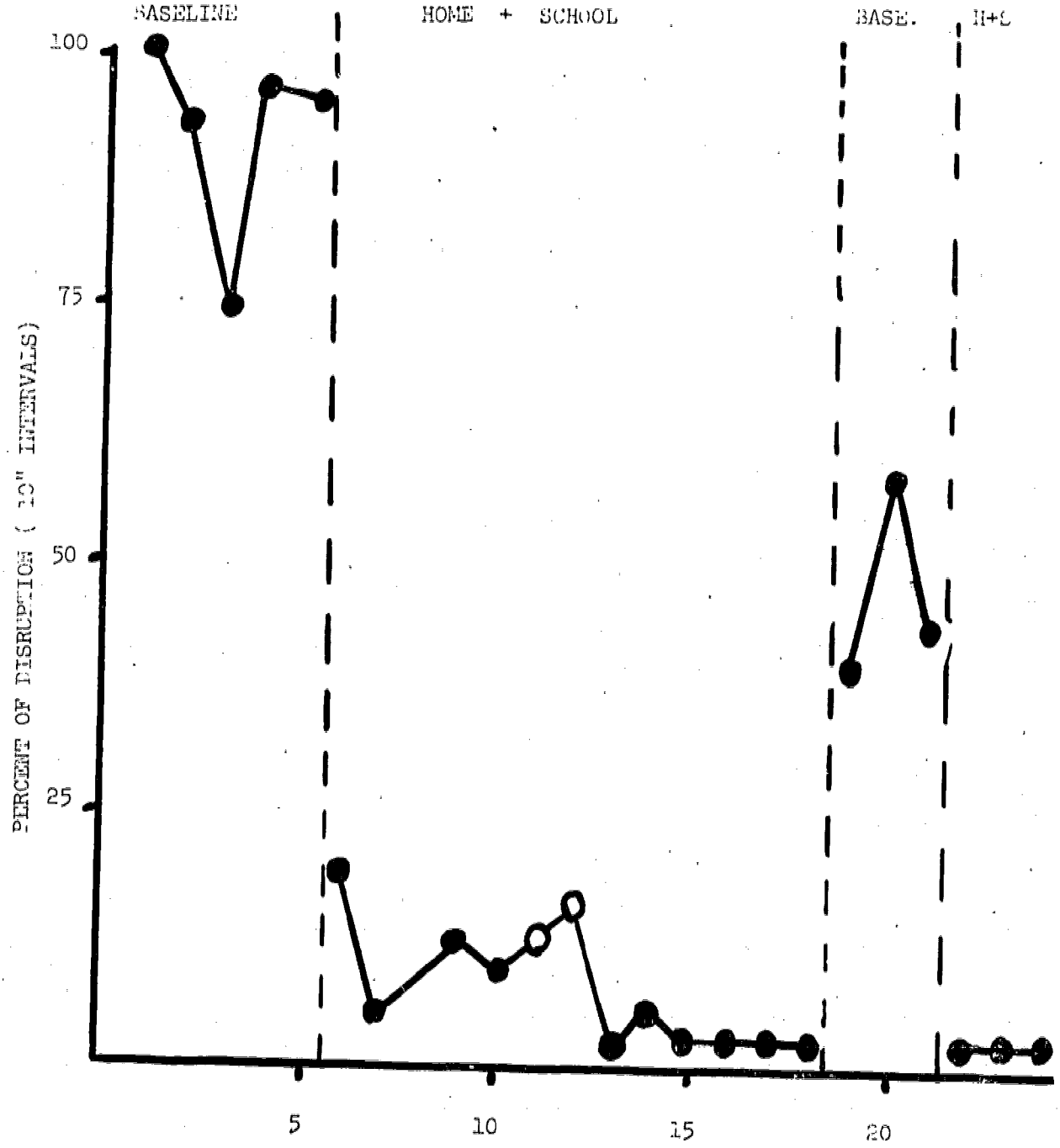
Results

Figure 3 reveals that during the implementation of the home-based management program disruption averaged less than 10 per cent. Prior to the behavior management efforts, disruption in this room had averaged between 90 and 100 per cent. When the "good letters" were given free in the morning on a noncontingent basis and staying after school was discontinued, disruption rose to around 50 per cent; but it immediately declined when the "good letters" were again made contingent upon good conduct and when the children were required to stay after school for misbehavior.

FIGURE 3

DISRUPTION- THIRD GRADE

N =23



Q__ SUBSTITUTE TEACHER

Discussion

Once disruptive behavior was effectively brought under control, it was possible for the teacher to concentrate her efforts upon the academic performance of the children. However, as only a very short period of time remained in the academic year (one and one-half months), a very intensive academic program had to be designed. The following experiment describes the program developed.

Experiment II -- The Effects of Multiple Ratio Schedules

Subjects

The subjects were the 23 children previously described in Experiment I.

Experimental Design

The experimental design employed for this experiment was an ABAB analysis in which a baseline was taken, the treatment procedure was applied, withdrawn, and reapplied.

Selection and Definition of Reinforcers

The reinforcers used were those already described in Experiment I.

Response Definition and Recording

The basic response measured in this experiment was the number of pages successfully completed in their reading workbooks daily. In order to successfully complete a page, a child had to get at least 70 per cent of the problems on that page correct. The teacher used, for the workbooks, the "quick" grading method developed in the second grade, except that she graded the pages herself.

Procedures

Prior to the beginning of the present experiment, the reading skill material had been presented in the traditional manner. Children were grouped

according to the level at which they were reading. These children were in levels 2 through 6 of the Houghton Mifflin series. Each day the teacher would work for 10-15 minutes with each reading group. She would explain and drill on material in the workbook and/or the actual reader for a particular level. While she taught one group, the other groups worked on an assignment of a limited number of pages in the workbooks.

The first step in introducing the new method for presenting material was to determine, for each child, whether he was currently placed at a level which was equal to his competence. Accordingly, the children were tested in the manner described in the "General Procedures" section.

The second step of the new presentation method was to eliminate any lessons in the workbook which the child had already learned. This was accomplished, again, by the level tests. Each test is divided into subtests which measure specific skills. When a child made a passing score on a subtest, the workbook pages concerned with that skill were deleted from his program of study. This process was repeated each time the student entered a new level. Once the baseline level had been gathered, the criteria for earning reinforcement was stated. In this experiment, the children were able to earn increasing numbers of points for increased amounts of work in what might be referred to as a multiple-ratio schedule of reinforcement.

A. Baseline. In order to control for the possibility that points per se, rather than points by multiple-ratio were responsible for performance increases, points were awarded in the baseline condition on a lean fixed ratio basis (3 points for each 2 workbook pages completed correctly). Points were not a complete novelty as they had been delivered contingent upon good behavior throughout the school year.

On the first day of the baseline condition each child was given a new workbook containing his individualized program of instruction (the list of pages he was to complete). The class was told that "reading" would now be done in a different way. They would be able to earn points for workbook pages completed correctly. In addition, there would be no more reading groups. The entire period (60 minutes)

would be spent on the workbooks. When they had questions, they were to raise their hands and come up to the teacher for explanation. A sign taped on the wall during this period gave the ratio of pages to points. The sign gave the current ratio in effect in each condition. The pages were grades and points delivered later the same day in time for Game Room.

B. Multiple Ratio. As the workbooks were handed out on the first day of this condition, the students were told that they would have an opportunity to make a great many points -- by doing a large number of pages in their workbooks. The ratios were as follows: for 2-3 pages correct, 3 points (same ratio as Condtion I); for 4-8 pages correct, 15 points; for more than 8 pages correct, 35 points.

C. Baseline. As the workbooks were passed out on the day beginning this condition, the announcement was made that the ratios prevailing in Condition I were again in effect, and the children were reminded what those ratios were.

D. Multiple Ratio. The multiple-ratio procedure was re-established during this phase, but as a systematic replication, one change was made in the point scale of Condition II. Instead of being paid 35 points for 9 or more pages correct, the students were paid 25 points. The purpose was to determine whether or not the children would work as hard on a learner reinforcement ratio.

Results

The class average of number of pages correctly completed per day is shown in Table 9. Following the institution of the Baseline condition, the average number of pages correct settled to a stable figure, between 2.4 and 3.0 with a mean of 2.65. After the institution of the multiple-ratio schedule, the number correct immediately rose to 5.2 and remained between 5.2 and 5.9 with a mean of 5.32. Upon the return to Baseline, the number dropped to 3.0 and stayed between 2.7 and 3.3 with a mean of 2.98. With the return to the multiple-ratio condition (with a reduced ratio), the number correct dropped to 4.0. It should be noted, however, that when observation was terminated

the curve had reached 4.5 and was rising. The mean during the last phase was 4.10.

TABLE 9

NUMBER OF CORRECT PAGES TURNED IN

Baseline	Multiple Ratio	Baseline	Multiple Ratio
2.65	5.32	2.98	4.10

Discussion

Although the multiple-ratio schedules generated more academic performance in a day-to-day basis, as measured by number of pages worked correctly, measures of the overall effectiveness of the program in generating increased academic achievement are still necessary. The following experiment describes an effort to achieve this goal.

Experiment III -- Levels Passed

Subjects

The subjects for this experiment included the 23 children described in Experiments I and II.

Response Definition and Recording

For this class, the amount of work necessary in order to pass a level varied slightly from that required for the other three classes in the program. Rather than working every page in their reading workbooks, the children were only required to work the pages that covered the areas their diagnostic test had shown their weaknesses. After finishing this work though, they were required to take and pass a test over all the work in that level before earning a trophy or being allowed to advance to the next level.

For example, if a child failed his level 4 placement diagnostic test and analysis of the results indicated that he failed because he was very weak in dictionary skills, he was only required to work the pages in his level 4 workbooks that dealt with dictionary skills. After successful completion of all those pages, he was allowed to take his level test.

The children only had 25 academic days during which they were able to pass levels for points and trophies.

Experimental Design

The experimental design for this experiment involved simple measurement of the number of levels passed during the period of reinforcement for reading performance in a pre-post analysis.

Selection and Definition of Reinforcers

The reinforcers utilized have already been described in Experiment I.

Procedure

The procedures for this experiment correspond exactly with those employed in the previous experiment.

Results

Table 10 shows the number of children passing reading levels in this class during the 25 days. An examination of this table will reveal that a total of 16 levels were passed by 12 children. Of the total of 23 children then, over one-half of them passed an entire reading level in only 25 academic days. This is a rate of 1 level passed every 1.56 academic days, the best rate attained by any of the experimental classes.

TABLE 10

READING LEVELS PASSED IN THE THIRD GRADE*

Number of Levels Passed	No. of Children
Number of children passing one level	9
Number of children passing two levels	2
Number of children passing three levels	1
Total Children	12
Total Number of Levels	16

* 25 academic days.

Discussion

The reinforcement program had success in generating both increased academic performance, as measured by number of pages worked correctly each day and increased academic achievement, as measured by number of levels passed. The statistical analysis of the Metropolitan Achievement Tests (MAT) scores would not provide any measure of the effectiveness of the program as both MAT tests (Fall and Spring) were administered before the program began. In an effort, though, to measure the amount of academic achievement present in this class prior to the implementation of the program for facilitating academics, a statistical analysis of these MAT scores was carried out.

Experiment IV -- The Statistical Analysis of the Metropolitan Achievement Tests

Subjects

The subjects for this experiment were the experimental group which included the 23 children already described and a control group which had 17 children between 8 and 11 years old. The mean age of the control group was 8.7 years.

Experimental Design

The experimental design for this experiment involved (1) a statistical analysis of the diffences in the gain achieved by the experimental and control groups from the Fall to the Spring administrations of the Metropolitan Achievement Tests (MAT) and (2) a statistical analysis of the mean score of the experimental group on the Fall test as compared to the mean score of the experimental group on the Spring test.

Procedure

Both Fall and Spring testings were carried out through the use of the standard testing procedures prescribed for administration of the MAT. The scores were then subjected to statistical analysis by means of a T test.

Results

The results of this experiment revealed (1) no statistically significant differences between the experimental and control groups and (2) no significan differences from Fall to Spring testing for the experimental group.

Discussion

The results of this experiment should not be unexpected. The experimental procedures for facilitation of academic performance were not implemented until after both testings on the MAT, thus, there should be no differences between the experimental and control groups. The failure of the experimental group to show significant gains when its Fall scores were compared to the Spring scores indicated that there had been little academic achievement in this class prior to the introduction of the behavior modification procedures; thus making the gains, in both day-to-day performance and overall achievement, attained with the behavior modification procedures even more impressive.

General Discussion

The results of the experiments conducted in the third grade pointed ou the value of linking the performance of children in school to some contingencies in their homes. The behavior management procedures were only minimally

effective until the children's parents were consulted and made a part of the program. The daily feedback upon the performance of their children kept the parents involved with the school and served to assist the school in managing the children. The systematic nature of the program resulted in very effective methods for maintaining orderly behavior in an extremely recalcitrant and disorderly group of children. Up until now, there had been a division of work, with the school being responsible for the academic performance of the children and for their conduct while they were in school and the home being responsible only for seeing to it that the children got to school. The results from this class made it very clear that there were ways in which the home could help the school to achieve its objectives. Parents had a great stake increasing their participation in the educational process of their children, as such participation could make the school much more likely to achieve its goal of quality education for every child.

Also, methods were developed, through the use of multiple-ratio schedules, for maximizing the amount of academic work done by the children. The use of the multiple-ratio schedules demonstrated that by offering the children larger amounts of reinforcement for increased amounts of work, the amount of correct work turned in could be increased when compared to offering the fixed amounts of reinforcement for each page turned in.

The fact that there were no significant differences between the MAT scores of the experimental group and the control group should have been anticipated. The program for maximizing academic performance was not implemented until after the Spring administration of the MAT; thus, the children were being taught without the benefit of reinforcement for academic points during the period from Fall to Spring MAT. The fact that the children in this class increased 16 levels in only 25 days following the implementation of the multiple-ratios program demonstrates the efficiency of such a program for improving academic performance.

SIXTH GRADE

Introduction

In the sixth grade, the major problems seemed to revolve around the relative inexperience of the teacher. As she was in her first year of teaching, her lack of experience led her to react in ways that were not always conducive to adequate control of discipline problems. Baseline observations indicated that while quantitatively the disruption in this class was not excessive (around 25 per cent), qualitatively the disruption was very intense. That is, while disruption did not occur as often as in some other classes (the third grade for example) when disruption occurred it invariably was very intense and resulted in a fight or in children being dismissed from the class. Consequently, the teacher indicated her desire to learn procedures which might lead her to manage more effectively this disruption rather than assisting her in academics. In any case, it is often better to start a teacher who has never been exposed to behavior management procedures with a relatively simple "quiet point" system rather than starting her in a full scale academic reinforcement system. Such a procedure allows the teacher to become accustomed to the mechanics of behavior management systems. Also, as disruptive behavior is so painfully obvious to a teacher, any system which helps her reduce the level of disruption will be accepted much more fully, thus, allowing her to implement academic reinforcement procedures with full confidence. Therefore, the first efforts in the sixth grade involved implementation of procedures for control of disruption. The following experiment describes these procedures and their results.

Experiment I -- The Control of Disruptive Behavior

Children

The children included all 31 students in this sixth grade class. All 31 children were black. Their ages ranged from 10 to 13 with a mean of 11.8 years. The IQ's of the children ranged from 61 to 134 with a mean of 80.85.

Response Definitions and Recording

Disruptive behavior was defined as:

A. Noise behavior: This included all noises made by the clas such as talking out, shouting, foot stomping, clapping, beating on desks, or call out the teacher's name.

B. Gross motor behavior: This included any motor behaviors such as out of seat, fighting, hitting each other, throwing objects, and kicking each other.

The observations were taken between 10:00 a.m. and 11:15 a.m. each morning for 27 academic days. The observer sat in the back of the room and observed the class by rows, as was done for the third grade. Each row was observed for 4 minutes each, then the observer shifted his observations to the next row. After observation of all 4 rows, the observer shifted back to the first row and began the cycle again. The 4 minute blocks were divided into 10-second intervals; thus, giving a total of 24 intervals for each 4 minutes during which a row could be marked as disruptive or non-disruptive. Each row was observed 4 separate times each day; thus, giving a total of 16 minutes or 384 intervals of observation per day.

Reliability checks were taken by a second observer on 6 of the 27 days. Interval by interval comparison of the data revealed that the reliability of the observations ranged between 89 per cent and 98 per cent with a mean of 94 per cent.

Selection and Definition of Reinforcers

As with the second and third grades, efforts were made to select reinforcers which occurred naturally in the school environment and were relatively inexpensive. Some reinforcers such as game room occurred only at set times each day while others like break could occur at any time during the day. Table 11 presents the list of backup reinforcers used in the sixth grade and their cost.

TABLE 11

LIST OF REINFORCERS AND THEIR PRICES
FOR THE SIXTH GRADE

Reinforcer	Cost
Recess	20 points
Break	5 points
Game Room	20 points
Basketball (rent)	3 points
Softball (rent)	5 points
Softball Bat (rent)	2 points
Softball Glove (rent)	2 points
Candy (in game room)	1-10 points
Chew Gum (all day)	5 points
Comics (buy)	5 points
Magazines (rent)	5 points
Play Radio (in game room)	5 points
Play Record Player (in game room)	5 points
Games (in game rom)	5 points

Experimental Design

The experimental design employed consisted of an ABA or reversal design. For this design, baseline observations were taken, the experimental treatment, the "quiet" point system was then applied and finally withdrawn.

Procedure

The class day was divided into 20 minute periods through the use of a kitchen timer, and the children earned "quiet" points for each 20 minute block during which they caused no conduct problems.

The procedure for reinforcing good conduct generally followed the procedure developed during the 1970-71 project at J. M. Jones School. Each child's name appeared on a lined mimeographed sheet which the teacher kept on a clipboard. The sheet was divided into squares representing the

20-iminute segments into which the day was divided. If the teacher saw a child engaging in disorderly behavior during a 20-minute period, she would announce "Howard, you are talking (fighting, playing, etc.) and you have lost your quiet point for this period. You can earn another one next period, though." This procedure was carried out for each child observed to be disruptive during a single 20-minute period. If the same child was observed engaging in disorderly behavior a second time during one 20-minute period, the teacher would say: "Howard, you are talking. That is the second time I have had to call you down so you lose all the quiet points you have earned so far today." This meant that Howard lost all his points which he had earned for good behavior. If a given child was called down a third time during a single 20-minute period, he then was required to write a sentence 500 times. The quiet points earned were marked on the children's cards twice a day, at 11:00 a.m. and 1:00 p.m.

All points earned had to be spent by Wednesday afternoon and Friday afternoon as the children were started out with zero points on Monday and Thursday mornings. Any points not spent by 3:00 p.m. on Wednesday and Friday were forfeited. The children could, however, save points from day to day during the other days of the week.

Results

The results for this experiment are presented in Table 12. During the initial 9 days of baseline, disruption averaged 23.11 per cent per day with a range from 60 per cent to 10 per cent. During the 9 days when good behavior was reinforced, disruption averaged 4.78 per cent with a range from 10 per cent to zero, and during the 9 days of return to baseline or no reinforcement, disruption averaged 23.44 per cent with a range from 35 per cent to 16 per cent.

TABLE 12

LEVELS OF DISRUPTION

	Level of Disruption (Per Cent)
Baseline	23.11
Behavioral Treatment	4.78
Baseline	23.44

Discussion

Once disruption was brought under control and the Behavior Research Staff felt that the teacher had become familiar with the use of behavior management systems, a system for facilitation of academic performance could be implemented. The following experiment describes the procedures and results of that experiment.

Experiment II -- Levels Passed

Children

The children included the 31 previously described sixth graders.

Response Definition and Recording

The level, as previously defined, consists of the material contained in one textbook. The teacher had, for each child, a sheet with the number of pages in his arithmetic and reading workbooks marked in it. As the child successfully completed a page, the teacher simply marked through that page on her sheet. This provided a constant record of where the child was working in each book, and when completed provided a record of the number of levels completed.

Selection and Definition of Reinforcers

The reinforcers utilized in this experiment included all the reinforcers previously described for the sixth grade in the previous experiment (see Table 9). Further, two reinforcers were added which were only available upon completion of a level. Table 13 presents the new reinforcers.

TABLE 13

NEW REINFORCERS FOR SIXTH GRADE

Reinforcer	Cost
Trophy	100 Points + Completion of a Level
Engraved Plate for Trophy	5 Points Per Letter

Experimental Design

The experimental design consisted of a simple measure of number of levels completed under the experimental condition, the reinforcement of successful academic performance.

The academic subjects for which the children could earn points included arithmetic and reading. When academic reinforcement began, the children were tested to determine the academic levels at which they should be placed. This procedure has been described in detail in the "General Procedures" section.

Procedures

The procedures begun in January involved giving the children one hour and fifteen minutes to work arithmetic and the same amount of time to work in reading each morning. During these periods the teacher would circulate about the room answering questions and helping children who were having problems. The arithmetic papers were collected at the end of the arithmetic period, and the reading papers were collected at the end of the reading period. When

both of these periods had expired, the teacher began to check the day's work, using the "quick" grading technique described for the second grade. The "quick" grading technique used in the sixth grade had some minor modifications, but functionally was identical with the techniques previously described.

Once all arithmetic and reading papers had been marked, the points earned by the children for that morning's work were placed on their point cards.

Concurrently, the "quiet point" system described in Experiment II was operating to help the teacher maintain an orderly class.

In arithmetic, the child was required to take chapter tests as he completed each chapter in the book. If a child failed this test, he was required to review the material in that chapter, with the teacher, and re-take the test. He had to pass the test before being allowed to proceed to the next chapter. Each page of arithmetic turned in that was completely right (100 per cent) earned the child 2 points and each page that was between 70 and 99 per cent correct earned the child 1 point. Scoring between 70 and 99 per cent on a "chapter" test earned 15 points and a score of 100 per cent earned 20 points.

In reading, as the workbooks were not divided into chapters, each 15 pages were considered a chapter. Therefore, a child had to take his chapter test after working 15 pages. The points earned for pages of work and for "chapter" tests corresponded with those earned in arithmetic.

In both subjects, once an entire book was finished the student was required to take a test over that book before being allowed to proceed to the next book. Successful completion (scoring $\geq$ 70 per cent) on this test earned the student 100 points and the right to buy a trophy.

Results

The results of this experiment are presented in Table 14. From this table it can be seen that 10 reading levels were passed in 100 academic days and 8 arithmetic levels were passed. Fourteen children accounted for the 18 levels that were passed -- one child passed one reading and two arithmetic levels and another child passed two reading and one arithemetic level. This is an overall rate of one level passed every 5.56 days.

TABLE 14

LEVELS PASSED (ARITHMETIC AND READING)
IN THE SIXTH GRADE

N = 31

Arithmetic	
Number of children passing one level	6
Number of children passing two levels	1
Total Number of Children Passing Levels	7
Total Number of Levels Passed	8
Reading	
Number of children passing one level	8
Number of children passing two levels	1
Total Number of Children Passing Levels	9
Total Number of Levels Passed	10
Number of Children Who Passed Both Reading and Arithmetic Levels	2
Total Number of Children Who Passed at Least One Level	14

Discussion

The results of this experiment indicated that in 100 academic days, approximately one-half of the class passed at least one level. These results indicate that there has been academic achievement in this class. In an effort to assess the significance of this academic achievement, another experiment, a statistical analysis of the results of Metropolitan Achievement Tests, was considered prudent.

Experiment III -- Statistical Analysis of the Metropolitan Achievement Tests

Subjects

The subjects for this analysis included the 31 previously described children as the experimental group plus a control class of 18 children. The ages of the control class ranged from 10 to 13 years with a mean of 11.3. Their IQ's ranged from 66 to 112 with a mean of 87.10.

Experimental Design

The experimental design for this experiment involved a statistical comparison between the gain manifested on the Metropolitan Achievement Tests (MAT) by the experimental group and the gain manifested by the control group on the same test. The independent variable was the presence of points for academic work in the experimental class, plus the presence of reinforcement for taking the test.

Procedure

The standard posttest on the MAT was administered by the Atlanta Public Schools in mid April. Comparisons between the amount of gain achieved by the experimental group as compared to the control group revealed no significant differences. As, however, there were approximately 30 academic days remaining in the school year, the MAT was administered for a third time during the last week of May. Further, for this testing the children in the experimental group were able to earn points for each MAT item that they answered correctly. These points were awarded immediately following the completion of each subtest. These points could then be exchanged for the regular backup reinforcers available during the course of the day. Other than this change, so that the children could earn points for their efforts, the MAT was administered in the standard fashion. For the control group, the MAT was administered in the standard fashion.

Results

Comparisons between the amount of gain from the Fall versus the May testing for the experimental and control groups revealed that for two of the subtests, reading and arithmetic concepts and problem solving, the experimental group had significantly more gain. For reading, the T value was 1.748 which was significant at the .05 level with 37 degrees of freedom. For arithmetic concepts and problem solving, the T value was 1.926 which was significant at the .05 level with 37 degrees of freedom. The significance of these two subtests indicated that the experimental treatment, reinforcement of correct work in reading and arithmetic, did in fact result in more "learning" as measured by the MAT.

General Discussion

The results obtained in the sixth grade indicated that (1) disruption was successfully brought under control within one day and was maintained at a very low level. These results were in agreement with those obtained in the previous program at Jessie Mae Jones School. Not only was the quantity of the disruption reduced, but the teacher felt that qualitatively the disruption was virtually eliminated; (2) academic gains were made as demonstrated by the fact that 18 entire levels were passed in 100 academic days; and (3) the academic gains that were made were reflected in the results of the MAT scores as the experimental group scored significantly higher on two of the critical subtests than did the control group.

SEVENTH GRADE

Introduction

In this class as in two of the other three project classes, the primary focus of the intervention program was upon facilitation of the children's academic performance with secondary focus upon control of disorderly behavior.

The teacher of the seventh grade was in her second year of teaching, and had been involved with the 1970-71 Behavior Modification Program at Jessie

Mae Jones School. Consequently, the Behavior Research Staff was able to implement new programs in her room without having to spend great amounts of time for teacher preparation. All new procedures were tried in this class first, before being implemented in the other classes where the teachers were not as familiar with behavior modification procedures. For example, the new grading procedures were developed in this class before being utilized in any of the other three project clases. The development of a general set of procedures in the seventh grade allowed the Behavior Research Staff to implement the grading procedure in the other classes with relatively few changes. The first experiment describes the grading procedures developed in the seventh grade.

Experiment I -- Grading Procedures

Two different grading procedures were developed in the seventh grade. They will be presented in the order in which they were developed.

Study A - "Quick Grading"

Subjects

The subjects in this class were 17 black children between 12 and 14 years of age, with a mean age of 12.8 years. The IQ's ranged from 53 to 114 with a mean of 72.11.

Response Definition and Recording

The response here was performance on the pages of written work turned in from the arithmetic and reading texts. A percentage correct for each page turned in was computed.

Experimental Design

The experimental design employed involved comparing the score on each page turned in after scoring through the use of two different grading techniques. This procedure permitted the computation of correlation coefficients for the two techniques.

Procedure

The procedure involved having the teacher select the one problem from each page of work turned in that she felt was most representative of that page. She then graded that problem and gave the student two points if he correctly completed it. When the child worked the problem incorrectly, the teacher selected another problem for grading: if that problem was worked correctly, the student received one point for the page. When the second problem was worked incorrectly, the student was considered to have failed the page; he was required to rework the page and resubmit it. When the page was resubmitted, the process of grading was repeated.

In order to check the reliability of this grading method, the Behavior Research Staff gathered all the papers after class for a period of two weeks. They then completely graded each page using the teacher's editions of the books as their guides. The pages graded by the staff were cast into one of two groups: any page on which the student had scored $\geq$ 70 per cent correct was considered passed and any page on which the child scored less than 70 per cent was considered failed. Over 600 pages were graded during this two week period.

For purposes of comparison, any page on which the child had received either one or two points from the teacher, who was using the "quick" grading techniques, was considered passed and any page on which the child received no points was consided failed. A page by page comparison was then possible with respect to whether or not the student had "passed" or "failed" the page with the two grading systems.

Results

The results of this experiment revealed that the correspondence between the two systems in identifying children who either passed or failed was .92 in arithmetic and .97 in reading. The overall level of correspondence was .95.

A comparison between the amount of time required for the teacher to grade the papers using the "complete" grading technique when contrasted with the "quick" grading techniques revealed that the time required for

the teacher to grade the papers with the "quick" technique was approximately 65 per cent less.

Study B - The Use of Student Graders

Subjects

The subjects for this experiment were the 17 previously described seventh grade children.

Response Definitions and Recording

The response examined in this study was scores on pages of work turned in by the students. The work examined came from both arithmetic and reading.

Experimental Design

The experimental design used involved obtaining correlations between the scores earned by the children under the two different grading techniques.

Procedure

Each day the teacher would choose three volunteers from the class to help her with the grading for that day. After the 2-1/2 hour period of work these children would go to the front table where the teacher had placed the teachers editions of the arithmetic and reading books. The rest of the clas would then turn in their work. Each of the graders would check the work by using the teacher's editions and on the top of the page write the number right on that page over the total number of problems on that page. The graders were required to sign their names to each page and they were not allowed to check their own work. The teacher made periodic spot checks of each grader's work in an effort to increase accuracy of grading.

After the graders had finished their work for the day, the papers were given to the teacher who placed the percentage score on them and wrote the number of points earned on the page.

For a period of two weeks the teacher graded all papers after class in an effort to determine the accuracy of the children's grading efforts. During these two weeks, over 900 pages were graded.

Further, a daily record was kept of the amount of time the teacher spent in grading activities.

Results

The results of this experiment revealed that the overall correlation between the teacher's complete grading and the grading of the tutors was 97.8 per cent. In arithmetic it was over 99 per cent and in reading it was 94 per cent.

As compared with the amount of time required for complete grading, this method saved approximately 85 per cent of the teacher's grading time; and as compared with the previously described "quick" grading technique, the use of student graders resulted in a savings of approximately 20 per cent.

Discussion

The results of the two studies presented in this experiment indicated that meaningful savings in teacher time can be achieved through the simple expedient of altering grading practices. Not only is the amount of time required for grading drastically reduced, but the children received much more immediate feedback on their efforts, as the new grading techniques permitted their work to be graded almost immediately. The fact that the new grading procedures permitted the teacher to get as much work from the class as possible, without having to worry about whether or not the work can be graded enabled her to permit the children to set their own limits as to how much work they would do and also permitted the use of such measures as number of pages turned in correctly in efforts to assess the value of the point system.

During the course of the experiment, it became very obvious that, given their choice, the children would choose to work many more pages in reading than in arithmetic. The reasons for this choice on the part

of the children were rather obscure, but the problem was real: The program design for this class required that the children be given freedom of choice, yet when given such choice the students worked virtually no arithmetic. In an effort to solve this problem, the following experiment was conducted.

Experiment II -- Altering Academic Preferences

Subjects

The subjects were the 17 previously described seventh grade children.

Experimental Design

The design used in this experiment was an ABA analysis.

Selection and Definition of Reinforcers

The reinforcers utilized in the seventh grade corresponded exactly in type and price with those used in the sixth grade. (See Table 11 for detailed list.)

Response Definitions and Recording

For this experiment, arithmetic was defined as worked turned in at $\geq$ 70 per cent correct from whatever arithmetic books the children currently were working in. Reading was defined as pages of work turned in at $\geq$ 70 per cent correct for the reading workbooks that the children currently were working in. The manner in which the teacher kept her records of what pages were worked correctly in each subject, the date on which they were worked, and the child who worked the pages was described in the previous experiment.

The definition of a page of work was given in the previous experiment.

Procedure

Following the "Placement Procedures" for academics which placed the children at their proper level in both reading and arithmetic, the children

were given a 2-1/2 hour period each morning when they could work on any subject matter of their choice for as long as they chose. For example, a child could choose to work on only arithmetic one day, and he might continue to work in arithmetic for a week or more. No pressure was exerted to get the children to choose one subject over another. During this time the teacher moved about the room helping individuals or groups with their work whenever they experienced difficulty.

After the 2-1/2 hour work period was up, the children turned in their academic work for grading. Following the completion of grading at around 12:00 noon, academic points were placed on the children's point cards.

A baseline was taken for three weeks under "normal" classroom conditions. During this time, the children received the same number of points for correct pages of work in both reading and arithmetic, one point for a score of 70-99 per cent and two points for a score of 100 per cent. Following this three-week baseline, the children were instructed that they would now be given four points for each page of arithmetic in which they scored between 70 and 99 per cent and six points for each arithmetic page on which they scored 100 per cent. The reading points remained the same as during baseline conditions, one and two points per page. After this period of greater magnitude of reinforcement for arithmetic had been in effect for three weeks, conditions reverted to those in effect during baseline. The children had the 2-1/2 hour work period each morning, during which they could work on whatever subject they chose for whatever part of the 2-1/2 hours they chose.

Results

Table 15 shows that increasing the number of points available for one subject (arithmetic) as opposed to another (reading) will lead the children to increase the amount of work they do in the first subject. In this case, the children increased from an average of approximately .98 pages per day per child to an average of approximately 2.4 pages per day per child in arithmetic. More importantly, once the children began to work more arithmetic, they continued to work in this subject, even when the pay off was no longer greater than the pay off for reading.

TABLE 15

NUMBER OF PAGES WORKED IN ARITHMETIC AND READING

Condition	Reading	Arithmetic	Total
When Performance in Reading and Arithmetic Resulted in Equal Number of Points	2.70	0.97	3.67
When Performance in Arithmetic Resulted in Greater Number of Points Than Did Performance in Reading	1.70	2.38	4.08
When Performance in Reading and Arithmetic Resulted in Equal Number of Points	1.85	2.01	3.86

Discussion

Once the children were working approximately equal numbers of pages in each subject matter, and were receiving equal point values for equal amounts of work in each subject, it was possible to consider other problems. Although the previous experiment had indicated that varying the number of points would lead the children to select different subject matters, it did not establish that the points did in fact result in increased academic performance. Consequently, an experimental demonstration of the control of the point system over the academic performance of the children was still lacking.

The next experiment describes an effort to establish empirically whether or not the point system control led the academic performance of the students.

Experiment III -- Rate of Daily Academic Work

Children

The subjects for this study included the 17 previously described children.

Response Definition and Recording

This study dealt with pages of correct academic work done. A page of academic work was defined as all the work included on any single numbered page in the reading workbook or the arithmetic textbook. Page 36 of the reading workbook, as an example, might have only 3 questions and page 37 might have 20 questions. For the purposes of this study, however, each was considered simply one page. In order to have a page considered correct, the student had to get at least 70 per cent of the problems or exercises on that page correct.

The teacher had for each student, in both arithmetic and reading, pages with columns of numbers corresponding to the number of pages in whatever book the child was currently working. If the child was working in the third grade arithmetic book, and that book had 255 pages, then for that child the teacher would have a sheet with the title of the book and the child's name and date he started the book at the top of the page. Under that would be columns of numbers starting with 1 and ending with 255. As the child correctly worked a page the teacher simply marked through the number(s) corresponding to the page(s) he had worked correctly and would put the date to the page. The teacher also recorded the dates of chapter tests on these sheets and the child's score on those tests.

Selection and Definition of Reinforcers

The reinforcers utilized in this class corresponded with those employed in the sixth grade. (See Table 11.)

Experimental Design

The experimental design for this study was an ABA or reversal analysis. The behavior and number of correct pages turned in, was measured under conditions of reinforcement, when the reinforcement was withdrawn, and when reinforcement was reinstated.

Procedure

During this experiment, the procedures corresponded basically with those described for the previous experiment. That is, the children had a 2-1/2 hour period each morning to work on subject matters of their choice with grading immediately following. The children earned one point for each page of reading or arithmetic worked between 70 and 99 per cent correct and 2 points for each page worked completely correct.

During the reinforcement period of this experiment, the children received points for each page they turned in at $\geq$ 70 per cent correct. They received one point for each page that was between 70 and 99 per cent correct and 2 points for each page that was 100 per cent correct. During no reinforcement the children were told that they would no longer be able to earn points for arithmetic and reading pages, but that they should continue working as their grades would be important in determining whether or not they passed for the quarter. Finally, when reinforcement was reinstated the procedures corresponded with those in effect during Baseline.

Results

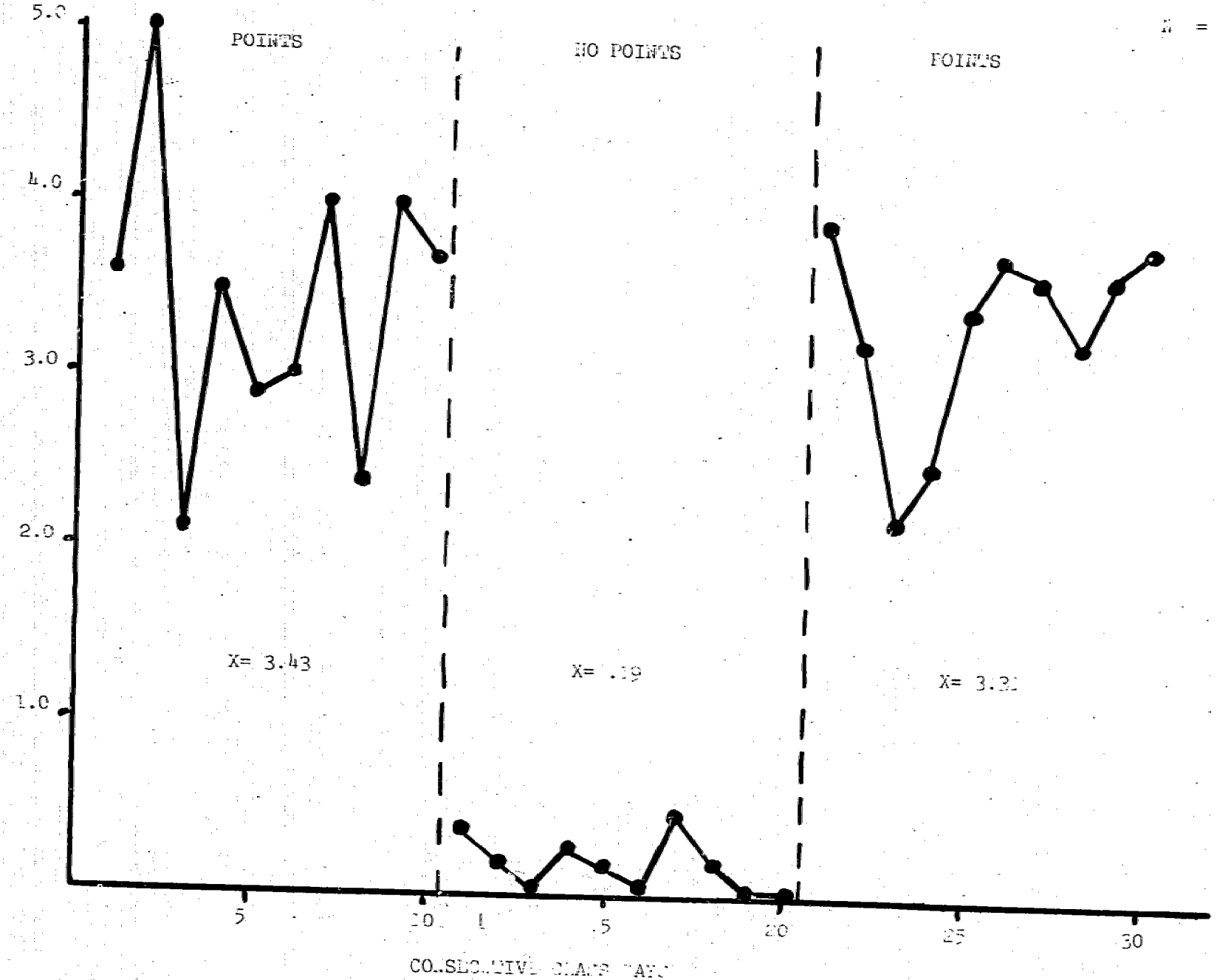
Figure 4 presents the results for this study. When the children could earn points for their work, they turned in an average of 3.43 correct pages per day; during no reinforcement, when no points were available for correct work, the children did an average of only 0.19 pages correct per day. When points were reinstated for correct academic performance, the average number of pages rose to 3.31, which was within 0.12 of its level under the previous conditions.

Discussion

The results of this experiment indicated that the point system was a powerful variable which strongly influenced the rate at which the children worked pages correctly in arithmetic and reading. It would be possible, however, for the children to work many pages each day in order to earn the points, and not integrate any of the separate things they learned.

FIGURE 4

MEAN NUMBER OF CORRECT PAGES PER DAY I.: ARITHMETIC AND READING



Chapters, however, typically required understanding of fundamental or central points which cut acorss individual pages. One method for determining whether or not the children integrated their knowledge would be to measure the rate at which they passed "chapter tests" under different conditions. The following experiment is an effort to establish experimentally the relationship between passing chapter tests in arithmetic and reading and the point system.

Experiment IV -- Number of Chapter Tests Passed

Children

The subjects for this experiment included the 17 previously described children.

Experiment design

The design for this study involved an ABA analysis. A baseline was taken of the number of arithmetic and reading tests taken over 10 days during which the children were reinforced with points for taking tests. Next the points for passing were withdrawn for a 10 day period and finally reinstated for another 10 day period.

Selection and Definition of Reinforcers

The children's reading books were divided so that every 15 pages was considered a chapter. The arithmetic books were already conveniently divided into chapters that averaged approximately 30 pages in length. When a child finished all the work in a chapter he was required to take a chapter test over the content of the material in that chapter. These tests typically consisted of 20-30 questions or problems drawn directly from the material. The teacher recorded the data and score of chapter tests in the sheets described in the previous study.

Procedures

The procedures used in this experiment corresponded with those used in the previous study with the following exceptions. The children were able

to take chapter tests any time they had: (1) finished all the pages in a chapter correctly and (2) requested a test from the teacher. If a child failed a test, the teacher reviewed the material with him and he was required to re-take the test. Any child who failed the same chapter test three times was required to rework all the pages in that chapter before being allowed to re-take the test. The chapter tests were taken at a special "test table" at the rear of the room where no talking was the dominant rule, and all chapter tests were taken during the 2-1/2 hours work period each morning.

During the first 9 days of reinforcement the children received 15 points for scoring between 70 and 99 per cent on chapter tests and 20 points for scoring 100 per cent. When the second 9 day period of no reinforcement was implemented, the children were told that they would no longer be able to earn points for chapter tests but that they were still expected to take and pass as many chapter tests as they possibly could. During the final 9 days, the reinforcement was reinstated.

Results

Table 16 shows that the children passed many more tests when reinforcement with points for test performance than when they were not reinforced. During the first 10 days, when they were able to earn points for their efforts on the tests, they passed 28 tests; whereas, when no points were available during the second 10 day period, they only passed 1 test. When the points were reinstated, the class took 27 tests in arithmetic and reading during the last 10 day block.

TABLE 16

NUMBER OF CHAPTER TESTS PASSED IN READING AND ARITHMETIC

(N = 17)

	Number of Chapter Tests Passed
Points for Passing Chapter Tests	28
No Points	1
Points for Passing Chapter Tests	27

Discussion

The results of the two preceding experiments indicated that the point system was effective in generating more academic behavior on both a day-to-day (pages) and chapter-to-chapter (tests) basis. These experiments did not, however, indicate the overall effectiveness of the point system in moving children to higher levels of academic achievement.

One measure of overall academic achievement was the level. To the extent that children were passing entire levels of work, it could safely be assumed that their overall academic achievement had been facilitated. Consequently, the next experiment attempts to measure the overall advancement of the children in arithmetic and reading.

Experiment V -- Levels Passed

Children

The subjects included the 17 children previously described plus one 13 year old black male who withdrew from school after the Christmas vacation.

Response Definition and Recording

The level has been previously defined and consequently will not be redefined here.

Selection and Definition of Reinforcers

The reinforcers used corresponded exactly with those presented in Tables 11 and 13 for the sixth grade.

Experimental Design

The experimental design consisted of simply counting the number of levels passed and the number of academic days during which the children could pass levels in a pre-post analysis.

Procedures

The procedures employed corresponded with those described for the previous experiments except that the children were required to take a test over all the material they had worked on when they completed a level. A score of 70 per cent entitled them to go on to the next level. If, however, they failed to score at least 70 per cent, the teacher reviewed the material with them and they were then required to retake the test. If any child failed to score at least 70 per cent on three successive attempts, he was required to rework all the material found on that level. There were 145 days during which the children could pass levels.

Results

Table 17 shows the number of children passing levels in both reading and arithmetic. This table indicates that of the 18 children involved in this experiment, 17 passed at least one level in either reading or arithmetic. Of the 17 children who passed at least one level, 13 passed more than one level, some children passing as many as 5 levels. Three of the 17 children passed levels in both reading and arithmetic. Thus, the table shows that while almost the entire class (17 out of 18) passed at least one level, many of the children passed more than one level, some passing as many as 5 levels.

TABLE 17

LEVELS PASSED IN READING AND ARITHMETIC
IN THE SEVENTH GRADE

N = 18

Arithmetic	
Number of Children Passing One Level	2
Number of Children Passing Two Levels	2
Number of Children Passing Three Levels	1
Total Number of Children Passing Levels	5
Total Number of Levels Passed	9

Reading

Number of Children Passing One Level	5
Number of Children Passing Two Levels	3
Number of Children Passing Three Levels	5
Number of Children Passing Four Levels	1
Number of Children Passing Five Levels	1
Total Number of Children Passing Levels	15
Total Number of Levels Passed	35
Number of Children Who Passed Both Arithmetic and Reading Levels	3
Total Number of Children Who Passed at Least One Level	17
Total Number of Levels Passed (Arithmetic and Reading)	44

Discussion

The experiments presented so far (except grading) have demonstrated that in arithmetic and reading, through the use of points for correct work, amount of daily work can be increased, the subject matter preferences of the children can be altered, the children can be motivated to pass more tests, and the children pass many entire levels. What has yet to be demonstrated, in this class or any other, is whether or not these procedures will increase the level or rate of academic performance in subjects other than arithmetic or reading. The next experiment was an attempt to answer this question with two new subject matters -- Science and Social Studies.

Experiment VI -- Science and Social Studies

Children

The children in this experiment were the 17 children utilized in the other seventh grade experiments.

Experimental Design

The experimental design consisted of a multiple baseline across two different subject matters, science and social studies.

Selection and Definition of Reinforcers

The reinforcers used were those already described in the previous sixth and seventh grade experiments.

Response Definitions and Recording

The two subject matters were taken from textbooks on the subjects. In social studies the book South America published by Harcourt, Brace, and World was used for all instruction and testing during this experiment and in science the books Prehistoric Animals and Our Solar System from the American Book Companies' Thinking Ahead in Science series was used.

Procedure

The children were instructed, as a group, in science or social studies sessions of the classes for approximately 30 minutes. Following the instruction period, the class was given 10 questions, multiple choice, true-false, and fill in the blank test over the material just covered. The teacher made up the tests and scored them immediately following their administration. During the Baseline [non-reinforcement periods], the class received no points for their scores on the tests, although they were reminded that these scores would make up their quarterly grades for these subjects. Next the children earned reinforcement in the form of points for passing their social studies tests, but no points for science. During this time, the children received 15 points for scores between 70 and 99 per cent and 20 points for scores of 100 per cent in social studies. Finally, the children continued to receive points as before in social studies, but were allowed to earn the same number of points in science. That is, for scoring between 70 and 99 per cent on science tests, they earned 15 points; and scores of 100 per cent earned 20 points. When reinforcement was available for science and/or social studies, the children were not allowed to earn points in arithmetic or reading.

This procedure was implemented in an effort to maximize the control of the points over performance in science and social studies.

Results

Table 18 presents the results of this experiment. This tables shows that when the children were reinforced with points for passing tests that their scores increased by at least 20 per cent over the previous level, when they could not earn reinforcement. These increases are statistically significant at the .05 level.

TABLE 18

AVERAGE SCORES ON SCIENCE AND SOCIAL STUDIES TESTS IN THE SEVENTH GRADE

N = 17

Condition	Science	Social Studies
No Reinforcement	54.2	61.6
Social Studies Reinforced Science Not Reinforced	54.9	81.7*
Both Reinforced	87.9*	85.3*

* Requirement for Reinforcement -- Score $\geq$ 70 per cent.

Discussion

The results of the experiments presented indicated unqualified success, and superiority for use of a point system as compared to a no point system. The results may not, however, extend beyond the boundaries of this class. That is, while the results indicated the efficiency of points on within group measures, it is possible that when compared with other classes through the use of standardized test scores it would become obvious that such a system was not necessary for academic achievement in other classes as measured by the standardized tests. The last experiment, between a group

comparison on the scores of both this class (the experimental class) and a preselected control class, was an effort to clarify this issue.

Experiment VII -- Statistical Analysis of the Metropolitan Achievement Tests

Subjects

The subjects in the experimental class included the 17 previously described children. The control class was another seventh grade from the same school. The mean age of the control group was 12.4 years and the mean IQ's for the control group was 86.20.

Experimental Design

The experimental design utilized for this experiment was a statistical analysis. The gain from Fall to Spring for the experimental group was compared with the gain for the control group during the same period of time. The independent variable was the presence of the reinforcement system in the experimental class.

Procedures

The scores from the standard Atlanta Public Schools' Fall and Spring Metropolitan Achievement Tests (MAT) were used in making these comparisons. During the period between testings, the experimental group was taught through the use of the point system, whereas, the control group was taught with standard classroom techniques. The teacher for the control group was very experienced with over 10 years experience, and she also had two Master's degrees.

Results

Comparing the mean gain scores for the two groups resulted in statistically significant differences in favor of the experimental group on reading [$T = 2.26$, $P > .025$, $(df = 31)$]; arithmetic concepts and problem solving [$T = 23.97$, $P > .001$, $(df = 31)$]; arithmetic computation [$T = 1.99$, $P > .05$,

(df = 31)]; social studies information [$\underline{T}$ = 1.99, P > .05, (df = 31)]; language [$\underline{T}$ = 1.896, P > .05, (df = 31)]; and overall MAT [$\underline{T}$ = 2.318, P > .025, (df = 31)].

It is of interest to note that the average IQ's of the experimental group was 71.11 while the average IQ's of the control group was 86.20. This difference is statistically significant. [$\underline{T}$ = 3.778, P > .001, (df = 30)]. As a consequence of these scores, an analysis of covariance currently is being conducted. The analysis of these results is not yet completed, but will be made available later.

Discussion

The experiments carried out in the seventh grade are critical to the whole 1971-72 program at Jessie Mae Jones School. It was in this class that new procedures were first implemented in an effort to iron out any major problems before beginning the new procedures in the other classrooms.

The results of the experiments which pertain to academic performance were extremely clear for this class. When allowed to earn points for their academic work as compared to when no points were available, the children turned in many more correct pages per day and passed more tests. Further, the number of levels passed and the significance of the experimental group as compared to the control group indicated that not only was the program successful over the short term (day-to-day), but it was very successful over the course of the academic year. Experimental control of academic performance then, has been demonstrated through the use of several different techniques and measures, the success of each giving more strength to the success of the others. Further, not only has the experimental control been demonstrated over arithmetic and reading, but also over science and social studies.

Taken as a whole, these experiments give strong evidence that a point reinforcement system can be an extremely effective procedure for (1) generating more day-to-day work and (2) bringing about significant long-term academic gains.

The experiment on altering academic preferences indicated that children can be encouraged to do the work that the teacher feels is most important by

simply altering the number of points available. For example, if she feels her class is not doing enough work in arithmetic, she simply increases the number of points available in arithmetic; thus, resulting in more arithmetic work for the class. Perhaps more important, the children will continue to work in arithmetic even after the points are returned to their previous values. This method offers a very efficient and pleasant way for facilitating work in the subjects that previously the children had found to be very unpleasant.

Finally, the experiments with grading techniques conducted in the seventh grade offered the teacher two methods for dramatically decreasing the amount of time required for scoring the children's work. The two methods could be used together, with the teacher using the grading technique that seemed most appropriate for any given subject matter or time, or the teacher could choose the grading method she liked best and use that technique almost exclusively. Either technique results in important time savings with little loss of accuracy.

In summary, the results of the experiments conducted in this class gave strong evidence for the value of a reinforcement system in facilitating academic performance in all subject areas and over long periods of time. Further, the children can be encouraged to engage in whatever academic performance the teacher feels is most appropriate through the use of differential point values. Also, grading can be dramatically speeded up through the use of simple grading techniques.

General Discussion

The results of this year's project at Jessie Mae Jones School indicated that a large scale program for modifying academic and disorderly behavior was implemented effectively over long periods of time. Many of the problems involved in utilizing behavior modification procedures for academic performance were solved through the use of either the "quick" grading methods or the student grading procedures.

The standard procedures for control of disruptive behavior were proven effective in three of the four classes, and in the fourth class the standard procedures were combined with a home based program and resulted in extremely

effective control of disruption. This home based management system provided the parents with daily feedback upon their children's conduct in school each day. The success of the home based procedures in controlling disruption points to the use of similar procedures for facilitation of academic performance.

A wide variety of academic behaviors were successfully dealt with during the year, ranging from arithmetic to social studies. The tactics for increasing academic performance were revolved around recognizing the children's efforts by allowing him to earn points for his work. In one class, tutors were used for helping the children and the results of this experiment indicated that not only do the tutors effectively help the children, but the tutors themselves receive significant academic benefits from such a procedure. The "natural" preferences of the children for one subject over another were easily altered through the use of differential point values, and the amount of work produced by the children was increased through the use of an increased number of points for increased amounts of work. The academic experiments also were effective in leading the children to pass large numbers of levels and resulted in significant gains on the Metropolitan Achievement Tests when compared to the control group.

To summarize briefly, an effective, long term wide scale program was developed -- one which effectively facilitates academic performance and controls disorderly behavior.

REFERENCES

Baer, Donald M., Wolf, Montrose M., and Risley, Todd R. "Some current dimensions of applied behavior analysis," Journal of Applied Behavior Analysis, 1968, I, 91-97.

Sidman, Murray. Tactics of Scientific Research. New York: Basic Books, 1960.