AN ABSTRACT OF THE THESIS OF

Abbas Hassan Ghandoura for the degree of Doctor of Philosophy in Science Education presented on February 22, 1982.

Title: Achievement Effects of Teacher Comments on Homework in Mathematics Classes in Saudi Arabia.

Abstract approved: Redacted for privacy
Howard L. Wilson

The purpose of this research was to study the effect of teachers' written comments on students' homework papers. Fourth, Sixth and ninth grade mathematics classes in Saudi Arabia were selected for the study.

The hypotheses for the study stated the following:

Teachers' comments significantly increase students' achievement in mathematics.

The effect of teachers' comments depend upon students' grade levels.

Six mathematics teachers in each grade level were randomly selected for the study in the city of Mecca, Saudi Arabia. Two classes for each teacher were randomly assigned, one as the Control group (No-Comment) and the other class as the Experimental group (Comment). For ten weeks the teachers collected homework papers three times a week. The homework papers in the Control group were graded by marking problems right or wrong. Homework

papers in the Experimental group were graded the same way, and had teachers' comments; i.e., comments--whatever the teacher chose to make.

An independent test measured whether these groups behaved differently from one another. Two forms of the final examination for each grade level were used as pre-post tests. A two-way analysis of variance was used to determine if students' achievement in the course was affected by the comments.

The results of the analysis of variance indicated that there was a significant difference in the mean achievement of the students in the experimental group. The treatment was effective in each grade level.

To test the relation between effectiveness of teachers' comments and students' ability, the students were divided into three different groups according to their performance in the previous term mathematics examination. All the groups responded equally to the treatment. The effectiveness of teachers' comments did not depend upon the ability of the students or grade level in the fourth, sixth and ninth grade mathematics classes.

Achievement Effects of Teacher Comments on Homework in Mathematics Classes in Saudi Arabia

by

Abbas Hassan Ghandoura

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1982

APPROVED: Redacted for privacy

Professor of Mathematics Education
in charge of major

Redacted for privacy

Chairman of the Department of Science and Mathematics Education

Redacted for privacy

Dean of the Graduate School

Date thesis is presented February 22, 1982.

Typed by Clara Homyer for Abbas Hassan Ghandoura.

ACKNOWLEDGMENTS

The amount of energy put into the successful completion of a dissertation can hardly be attributed to a solo effort. In the light of this, I would like to express candid gratitude to my advisor, Dr. Howard Wilson for the consideration and assistance he gave me from the comceptualization to the completion of this study. His helpful suggestions have been evidenced by his cooperation and continual encouragement.

Sincere appreciation is extended to the members of my doctoral committee--Drs. Thomas Evans, Karl Nice, Gary Musser, Joseph Hlebichuk, and Howard Wilson for their critical comments and timely encouragement throughout the stages of this study.

Deep appreciation is extended to Drs. Gene Craven, Fred Fox and to the entire professors in the Department of Science Education: Thomas Evans, Karl Nice and Howard Wilson. Their professional guidance will always be treasured. My interactions with these fine professionals have been intellectually rewarding.

Recognition is extended to:

Dr. Maggie Moore for her friendship and encouragement.

Dr. Dave Niess for his assistance in the computer analysis of data in this study.

Cathy Hall for her assistance in the design of the study.

My gratitude is extended to Dr. Courtney for his assistance in the experimental design.

I am especially grateful to my wife Susan for her steady encouragement, and patience throughout the preparation of this dissertation; and to my child Waleed for his understanding when his father was busy.

TABLE OF CONTENTS

Chapter		Page
I	INTRODUCTION	1
	The Problem	2
	Statement of the Major Hypotheses	3
	Statement of the Minor Hypotheses	4
	Need for the Study	7
	Difficulties of Educational Generalization	7
	Differences of the Usual Experimental Condition from the Classroom	9
	Assumptions of the Study	11
	Definition of Terms	11
	Limitations	12
	Design of the Study	12
	The Experimental Procedure	13
	Organization of the Remainder of the Study	15
II	LITERATURE REVIEW	17
III	THE STUDY	23
	The Purpose of the Study	33
	The Major Hypotheses	33
	The Minor Hypotheses	34
	The Experimental Procedure	37
	Research Design	46
	Statistical Analysis of the Data	50
IV	RESULTS OF THE STUDY	53
	Findings Related to the Major Hypotheses	53
	Findings Related to the Minor Hypotheses	58
	Group 1	58
	Group 2	62
	Group 3	67
	Summary of the Results	72
V	SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	76
	Summary of the Study	76
	The Problem	76
	Design of the Study	77
	Result of the Data Analysis	79

Chapter		Page
	Conclusions	80
	Recommendations	83
	Recommendations for Further Research	83
	Recommendations for Educational Practice in Saudi Arabia	84
BIBLIOGRAPHY		86
APPENDICES		
A	Recruitment Letter to Teachers	89
B	Analysis of Variance and Detailed Description of the Populations	97
C	Interaction	122
D	Sample of Teachers' Comments	127
E	Teacher-made Tests	129

LIST OF FIGURES

Figure		Page
1	Sample Selection Procedure	39
2	Interaction Between the Treatment Effect and the Grade for the Entire Population	57
3	Interaction Between the Treatment Effect and the Grade for Group 1	62
4	Interaction Between the Treatment Effect and the Grade for Group 2	67
5	Interaction Between the Treatment Effect and the Grade for Group 3	72

LIST OF TABLES

Table		Page
1	Number of Students in Each Grade Level	41
2	Summary of the Population in Group 1	41
3	Summary of the Population in Group 2	42
4	Summary of the Population in Group 3	42
5	Design-Matrix	49
6	Analysis of Variance	51
7	Analysis of Variance for the Major Hypotheses	54
8	Mean Gain in Achievement by Grades	54
9	Analysis of Variance for the Minor Hypothesis in Group 1	59
10	Mean Average Gain in Achievement by Grade for Group 1	61
11	Analysis of Variance for Minor Hypothesis in Group 2	64
12	Mean Average Gain in Achievement by Grade for Group 2	65
13	Analysis of Variance for the Minor Hypotheses in Group 3	69
14	Mean Average Gain in Achievement by Grade for Group 3	69

Achievement Effects of Teacher Comments on Homework in Mathematics Classes in Saudi Arabia

CHAPTER I

INTRODUCTION

Some of the most vexing problems before teachers and parents center about homework. Shall we have homework, or shall we not have homework? How long should pupils be expected to study at home? At what age, grade and school level should home study begin? In how many and in what subjects should secondary pupils be required to study at home? Can pupils succeed in school work without homework? Is it worth a teacher's time to inspect homework and write comments on students' homework papers? It would be highly desirable to answer all these questions for all subjects and all levels of school organization, but obviously that task is too comprehensive for one study. It must be attacked by many workers in each level of school organization. The aim of this study is to throw light on the general problem at one particular spot, and it remains for others to illuminate at other points.

There has been much discussion about homework. Frequently the principals of schools in the same neighborhood have different views about the matter, and as it often happens, especially when a principal is

careless or inefficient, that teachers in the same school are permitted to exercise varying policies in the assignment of homework.

A careful investigation of the rules adopted in various places for the regulation of homework shows a wide divergence of practice. In some cities the teachers are given full freedom as to the kind and amount of homework which may be assigned. It is extremely difficult for the teachers to give more attention to the needs of individual pupils especially in large classes.

THE PROBLEM

Each year mathematics teachers spend many hours writing comments on homework papers being returned to students. Apparently they believe that their words will produce some results in students' performance, superior to that obtained without comments. Yet on this point solid experimental evidence, obtained under classroom conditions, has been absent. One will comment copiously, another not at all, and each believes himself to be right.

The study is aimed at providing experimental evidence about the effects of teachers' comments on student performance. It investigates the question: Do teacher comments on students' homework in mathematics effect achievement in mathematics? It is also the aim

of this study to examine the relationship between the comment effect if any and students' ages or grade levels. Such a question will find its clearest answer in an experimental study; one in which some students receive teacher comments, and others receive no comments.

STATEMENT OF THE MAJOR HYPOTHESES

The major hypotheses are as follows:

Treatment

H_1: There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average score on teacher made tests.

Grade

H_2: There is no significant difference in the mean grade levels in fourth, sixth and ninth grade mathematics students, between the experimental group and the control group as measured by the average score on teacher made tests.

Interaction

H_3: There is no significant interaction between the treatment and grade levels of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured

by the average score on teacher made tests.

STATEMENT OF THE MINOR HYPOTHESES

In order to test the treatment effect on students' ability levels the investigator obtained the previous term mathematics score (out of 100 points) for each of the students. Three minor hypotheses for each main hypothesis were tested. The students were divided into three groups to test for which group the treatment worked best. These groups were:

Group 1. The students who scored up to 55 points on the previous term mathematics examination.

Group 2. The students who scored between 55 and 75 points on the previous term mathematics examination.

Group 3. The students who scored more than 75 points on the previous term mathematics examination.

The minor hypotheses are as follows:

Treatment

$H_{1.1}$: There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students (Group 1) in the experimental

group and the control group as measured by the average score on teacher made tests.

$H_{1.2}$: There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students (Group 2) in the experimental group and the control group as measured by the score on teacher made tests.

$H_{1.3}$: There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students (Group 3) in the experimental group and the control group as measured by the average score on teacher made tests.

Grade

$H_{2.1}$: There is no significant difference in the mean grade levels of fourth, sixth and ninth grade mathematics students (Group 1) in the experimental group and the control group as measured by the average score on teacher made tests.

$H_{2.2}$: There is no significant difference in the mean grade levels of fourth, sixth and ninth grade mathematics students (Group 2) in the experimental group and the control group as measured by the average score on teacher made tests.

$H_{2.3}$: There is no significant difference in the mean

grade levels of fourth, sixth and ninth grade mathematics students (Group 3) in the experimental group and the control group as measured by the average score on teacher made tests.

Interaction

$H_{3.1}$: There is no significant interaction between the treatment and the grade levels of fourth, sixth and ninth grade mathematics students (Group 1) in the experimental group and the control group as measured by the average score on teacher made tests.

$H_{3.2}$: There is no significant interaction between the treatment and the grade levels of fourth, sixth and ninth grade mathematics students (Group 2) in the experimental group and the control group as measured by the average score on teacher made tests.

$H_{3.3}$: There is no significant interaction between the treatment and the grade levels of fourth, sixth and ninth grade mathematics students (Group 3) in the experimental group and the control group as measured by the average score on teacher made tests.

NEED FOR THE STUDY

Educators usually all recognize the desirability of effective teacher-produced motivation, teacher behaviors which will in some way cause desirable student behavior. Yet the empirical literature bearing on teacher comments on homework is remarkably scanty. The necessity of the present study seems urged by at least three conditions which may be briefly demonstrated. First, teachers differ so widely in their commenting practices that their divergence indicates a lack of authoritative information on the subject. Second, it is very difficult to generalize securely from laboratory experiments to the ordinary classroom conditions. And, finally, experiments investigating matters related to teacher comments have done so under conditions varying greatly from true classroom conditions and true teacher-pupil relationships.

DIFFICULTIES OF EDUCATIONAL GENERALIZATION

As already noted, there is no appreciable literature directly concerned with teacher comments on students' homework papers. Rather there is a large amount of literature in the field of general psychology dealing with matters somewhat related to such comments. Of the rather few studies actually concerned with comments to

school children, none has been conducted by a random sample of teachers under strictly ordinary classroom conditions.

If school comments work at all, they might be considered analogous to what is called, within learning theory, 'reinforcement' in the usual sense (Skinner 1953). That is, these comments would operate by affecting the students' responses in the same general situation. If comments caused high achievers on one test, for example, to remain high and to avoid the usual regression on the next test, such comments might be called "positive reinforcement." If comments caused high achievers to drop below the regression point of those high achievers who receive no comment, then the comments might be called "negative reinforcement," and so on. One might question, however, which term should be employed if a comment caused a significantly <u>improved</u> performance in low achievers? Or to ask an equally puzzling question, which concept would be more appropriate to the "B" students; for example, who receive mild praise for the "B" and as a result significantly improve their status? Have the comments then acted as "positive" or "negative" reinforcers? In other words, under which theoretical concepts should one look for a prediction, or even an explanation of comments effectiveness? This definitional problem would appear important for the generalization of

psychological experimentation to classroom method, yet it has not been reconciled in the literature.

Brenner (1934) has pointed out the increasing desirability of research which is school-related. Among what he terms "weakness" of current studies certain incentives, which, in a certain way in a laboratory situation, will act similarly in every school situation.

Apparently then, there is ample evidence for supposing that educational research must not rely too heavily upon generalizations from areas of laboratory experimentation, except where laboratory conditions are similar to those of the classroom.

DIFFERENCES OF THE USUAL EXPERIMENTAL CONDITION FROM THE CLASSROOM

Once having admitted that the verbal label used to describe an event (such as "reward," "punishment," "praise" etc.) is not necessarily isomorphic to the event itself, one may reach certain conclusions about educational experimentation. From an educational viewpoint, research must be regarded as only tentative when its condition differs too greatly from those of the classroom. In order to indicate the importance of the present experiment, it may be desirable to inspect some conceivably related studies, noting any ways in which their conditions fail to be educationally realistic.

However useful to learning theory (and such usefulness is not here called in question), previous studies related to the present one have been encumbered by certain weaknesses. To cite one obvious example, studies have frequently been conducted with animals rather than humans. The general value of animal study is widely acknowledged. But it will probably not be disputed that, if any single assumption might cause generalization to be insecure, it might be the bridge from the nervous system of one species to that of a very different species. For the educator, animal studies must probably function primarily as heuristic models, not so much to answer educational questions, as to ask them.

In a similar manner, whether reinforcements were administered to animals or humans, such reinforcements have often consisted of shock, or some other incentive totally unused in the classroom. Such differences make the task of educational generalization all the more complex. It is not denied that such studies may have some relevance to a problem in teacher comment, yet it appears almost impossible to know precisely what the relevance might be without actually testing such relevance within the classroom situation.

The present experiment was designed to keep the total classroom procedures exactly as they would have

been without the experiment, except for the written documents themselves.

ASSUMPTIONS OF THE STUDY

In order to conduct this study, these assumptions must be made:

1. There is no discussion between the students in the experimental group and the students in the control group on homework assignments.

2. The teachers will follow the instructions concerning comments on mathematics homework.

3. The classroom conditions surrounding the experimental and the control classes are identical.

4. The length and difficulty of homework assigned by different teachers at a particular grade level is nearly enough alike to be considered the same.

5. Teachers' comments effect students' performance.

DEFINITION OF TERMS

Terms whose meanings are somewhat unique to this research study are those defined below.

1. Achievement --- refers to the acquisition of mathematically-treated knowledge as measured by the average gain on teacher-made pre and post tests administered before and after the experimental period.

2. Homework --- refers to written mathematics assignments of the traditional kind assigned by mathematics teachers, done by the student at home, and checked by teachers.

3. Interaction --- refers to the joint effect of the two independent variables on the dependent variable. Examples of interaction are found in Appendix C.

LIMITATIONS

This study is limited:

1. To fourth, sixth and ninth grade levels in Saudi Arabia.

2. To boys only enrolled in these grade levels.

3. To one subject matter only: mathematics.

4. By the validity and reliability of the teacher-made tests.

5. By teacher's ability to write appropriate comments on students' mathematics homework papers.

DESIGN OF THE STUDY

The experimental design selected for the study is the nested split-plot. The term split-plot comes from agricultural experimentation in which a single level of one treatment is applied to a relatively large plot of

ground (the whole plot), but all levels of a second treatment are applied to sub-plots within the whole plot.

In this experiment, levels of one factor (teachers) occur only within the level of another factor (grade)--thus, we might compare three different grade levels, and another factor of interest might be the teacher in a particular grade. Here the factor of "teacher" is said to be nested with the factor of "grade."

THE EXPERIMENTAL PROCEDURE

The steps in conducting this experiment are summarized as follows:

1. List all fourth grade mathematics teachers in the city of Mecca with two or more mathematics classes. From the list randomly select six teachers.

2. Repeat for sixth and ninth grade mathematics teachers.

3. Randomly select two classes for each teacher; one class as the experimental group and the other class as control group.

4. Give pretest on material to be taught during the study.

5. Apply the treatment.

6. Give post-test, which is a different form of the pretest.

7. Determine the difference between post-test score and pre-test score.

8. Use a two-way analysis of variance to test the hypotheses.

Thus the sample will consist of 12 classes in each grade level: six classes assigned as the experimental group and the other six classes assigned as the control group.

For ten weeks the teachers will collect homework three times weekly. The homework papers in the no-comment group (the control group) will be graded by simply marking problems right or wrong and giving the papers a grade. Homework papers in the comments group (experimental group) will also be graded by marking the problems right or wrong and giving the papers a grade. However, papers in the comments group will have teacher comments on students' homework. If a student missed a problem the teachers will mark where the error occurred, and either give or indicate how to do a correct solution. If a student does not do as well as the teacher expects, the student will be encouraged to do better next time.

Each paper in this group will have one or more written comments, free comments--whatever the teacher chooses to make. The teachers will return the homework papers in the customary way. If the teacher usually talks

about the papers being returned, he will be expected to do so again; otherwise he will not. It is important that the students remain naive about the experiment.

The teacher will go on with what the class would normally be studying. By the end of the tenth week the students will have a test. All the teachers who teach a particular grade level will participate in preparing the test. It will be objective; i.e., produce numerical scores without any individual decision about each paper.

The teachers will report each student's score on the paper and return the papers tc the investigator. Then from the teachers' point of view, the experiment will be over.

The outlined procedures permit certain experimental virtues: they permit a random sample of schools, teachers and classes. They permit the experimental treatment to be assigned under truly random methods standardized for all participant groups. They permit the classroom procedures, with the exception of one independent variable of the comments themselves, to be totally undisturbed, and permit a psychological treatment of many students without any of them even knowing that they are experimental subjects.

ORGANIZATION OF THE REMAINDER OF THE STUDY

The remainder of the study is separated into four

chapters. Chapter II reviews the studies involving homework. Chapter III details the design of the study. Chapter IV and Chapter V present and discuss the results of the study. Chapter IV analyzes the data and discusses the results relative to the hypotheses. Results not directly related to the hypotheses are included. Chapter V contains a summary, conclusions, and recommendations for further study.

CHAPTER II

LITERATURE REVIEW

Questions relating to homework have long been discussed in educational literature. From 1900 to 1974 there are approximately 400 articles on homework as home study referenced in the Educational Index. However, most of these articles are based on opinion rather than empirical evidence. In 1961 the editors of the National Educational Association Journal expressed the intention to publish an article about what research says about homework. They found that

> Very few experimental studies on the problem of homework and many of those reported are limited in number, scope, and quantity. For the main part, these studies are inadequate in design and do not involve carefully controlled and well-planned experiments in which the range of factors which influence learning are explored.

Two earlier articles reviewed the research on homework. Swenson, Cost and Taylor (1955) in summarizing the research before 1950 stated that

> While an occasional study ... seemed to indicate slight advantages associated with home study, most of the research before the stated date has been unfavorable to homework.

They reviewed four reports on homework problems published

in the 1950's. Two of the studies were surveys of opinions toward homework; a third study related how faculty, pupils and parents worked on the problem of home study. Only one of the four studies was an attempt to determine the effect of homework on student achievement. In this study they found that

> In the field of high school social studies it makes no difference whether homework is assigned or not, either in social studies abilities or in achievement test scores.

Swenson, Cost and Taylor concluded by stating that

> Whether one considers the older studies (before 1950) or the four more recent studies--it is apparent that additional research evidence is needed on the whole question of homework.

Goldstein conducted a comprehensive review of research on homework. His purpose was to examine the value of homework in the light of the pertinent research published during the thirty years before December, 1958. For this period the Education Index lists 280 titles on home study.

Of the 280 titles, only seventeen are actual reports of experimental research on the homework problem. None of these pertain to Grade 1 through 4; seven relate to Grades 5 and 6. Eight (including four of the previous seven) deal with Grade 7 through 9; and six concern Grade 10 through 12 (1960).

Goldstein abstracted and critiqued each of the seventeen reports of experimental research on homework. Five of the reports involved homework in arithmetic at the junior high school level. A brief report of each of the five studies as reported by Goldstein is given below.

> Teacher set-up matched groups of sixth-seventh and eighth graders for an arithmetic study in a New Jersey school. The experiment is only sketchingly described, and the design details are left unclear. Homework and non-homework groups were tested after 115 days. The median scores were the same in both groups. The author concluded that homework is of no value in promoting achievement in arithmetic.
>
> - - - - -
>
> Stenines examined a single seventh-grade class in Ingram, Pennsylvania. Half the pupils had compulsory homework in arithmetic but not in English; for the other half the assignments were reversed. Thus all thirty-nine pupils had half an hour of assigned homework daily, but in different subjects. The pupils in the two groups were roughly matched initially into pairs based on mental ability and achievement

in English and arithmetic. The test scores at the end of the semester showed clear gain in arithmetic achievement by the group that had assigned homework over the group that did not ...

- - - - -

Montgomery conducted a similar experiment with thirty pairs of pupils in Grades 7 through 9 in a West Virginia mining community of low economic status. The design differed only slightly from that of the study just described. One group was given homework in English but no homework in arithmetic. The other group was given homework in arithmetic but no homework in English. After one term the groups were crossed over and the experiment was continued until the end of the year. The mean intelligence quotient of the entire group was only 87. The overall data showed a slight advantage for homework in arithmetic--although the result was not statistically significant. The ninth grade and the groups with highest intelligence showed "a decided result in favor of home

study in arithmetic."

- - - - -

Foran and Weber conducted an experiment in seven parochial schools with 292 pupils in seventh grade. No homework was given in any subject except arithmetic and then only during one of the two terms under study. A cross-over design was used. Group A had homework during the second term, Group B only during the first. Both groups made greater achievement gain in problem-solving with homework than without. In computation the outcome was less clear but showed the same tendency ...

- - - - -

Anderson studied eighth-grade classes in an Oklahoma junior high school. Achievement gains in English, social studies, and mathematics were compared in classes that did and did not have regular homework assignments. Pupils were matched on the basis of mental ability. All classes in a given subject had the same teacher. Achievement was tested after each unit of work. The group with regular homework assignment showed higher achievement gains in all three subjects (1946; 217-9).

Goldstein concluded that experimental findings do

not warrant statements that homework contributes little or nothing to immediate academic achievement.

Hines (1957) studied the effect of homework on achievement in plane geometry in two high school classes in Urbana, Illinois. Nineteen pairs of students were initially matched by chronological age, I.Q. and previous performance in algebra. Members of each pair were randomly assigned to either the no-home-study group or the home-study group. Standard achievement tests in algebra and geometry showed that the two groups were similar on these tests. The same textbook was used for the two groups and the experiment lasted the entire year. However, each group was taught by a different teacher. Evaluation was restricted to achievement in plane geometry as measured by scores on eight unit tests, seven cumulative review tests, a semester examination and a standardized geometry test. Everyone of the seventeen comparisons favored the home-study group. Nine of the differences were significant at or beyond the five percent level. Hines' results can be questioned on two points: (1) The groups were small. The final results used only sixteen of the assigned nineteen matched pairs. (2) Each group was taught by different teachers. Hines stated that

> It is possible that the differences found could be accounted for by differences in skill

or zeal on the part of the two teachers involved.

Koch (1956) conducted a study with three sixth-grade classes in the same school. Each class was taught by the regular classroom teacher using the teacher's usual method of instruction and the same arithmetic textbook. During the ten-week duration of the experiment one class received one or two daily homework assignments of approximately 30 minutes, one class received short daily homework assignments of approximately 15 minutes, and one class received no homework. Koch's results indicated that homework significantly increased achievement in computation for the class doing the long homework assignment. As in Hines' study, Koch had the limitations of a small population (a total N of 85) and uncontrolled teacher variables.

Hudson's study investigated the relationship between the amount of assigned homework and its effect upon scholastic achievement in seventh grade mathematics. He found that the amount of homework assigned has no statistically significant relationship to scholastic achievement in arithmetic concepts. Scholastic achievement in arithmetic problem-solving may be influenced by the amount of homework assigned.

Mason (1965) attempted to determine whether required homework results in greater achievement. He had a total N of 432 students; 241 in nine classes were

required to do homework and 141 in nine classes were not required to do homework. Each teacher taught both types of classes. There was no significant difference in the results obtained by requiring homework or not requiring homework when averaged over all teachers.

> However, the significant level of teacher by method interaction was 99.5%. This high significance level leads to the conclusion that each teacher should determine which of the two methods to use in order to be most effective as a teacher.

This result seems to emphasize the earlier questioning of the results of Hines and Koch.

In 1964, Ten Brinke reported on his investigation of the effect of homework on achievement in mathematics in Grades 7 and 8. The contrasted treatments were homework and supervised study. Two classes of 27 students in seventh grade and two classes of 27 students in eighth grade were used. One class in each grade received daily assignments that were corrected and collected at the beginning of the succeeding period. The other classes received somewhat shorter daily assignments about midway in each class period. Significantly superior achievement relative to homework in contrast to supervised study was not found for classes taken as a whole. Some evidence indicated a differential effect of homework toward superior achievement for upper ability students and of

supervised study toward superior achievement for low ability students.

Studies by Bradley, Whelan, Maertens and Leonard were all conducted at the elementary school level. Leonard's study was in the area of fifth grade social studies. She found that homework increased student achievement in social studies. Bradley compared two methods of assigning homework in fifth grade mathematics. His results revealed significant differences favoring individualized homework assignments compared to the blanket-type homework assignments. Whelan investigated the effect of systematic homework assignments in English and arithmetic on the achievement of fourth grade pupils. He found no significant differences. Maertens examined the effect of arithmetic homework upon the arithmetic achievement and attitude development of third grade pupils. He found that arithmetic homework did not have a significant effect upon the arithmetic achievement of third grade students.

Zastsow conducted his study with college classes in calculus and classes in statistics. He studied the "effect repetition (in the form of solving homework problems) had upon the learning processes if the subject matter were meaningfully presented." The results of the study:

> Hinged upon presenting the material to the students in a meaningful manner and after surveying existing methods of instruction, a new one was devised for this project.

He stated that any meaningful method of instruction would have to make the student become an active participant. A four-step process to force student participation was outlined. How Zastsow determined that his method of instruction was meaningful is not clear. His results revealed significant differences favoring homework for acquiring skill in problem-solving for the calculus classes.

With the exception of Bradley all of the previous studies attempted to determine whether homework was better than no homework. Bradley attempted to determine if one method of presenting homework was better than another method. The only other study which attempted to find an improved method of assigning homework was Nietling's.

Nietling gave "problems to initiate the study of certain topics in mathematics" to college students in his classes. After a problem was assigned, the student had at least two days to formulate his conclusions. Problems were normally presented without any classroom introduction unless "the nature of the problem seemed to suggest that the typical student would not get started without some introduction" or "the solution of the problem required some information which the student had not yet studied."

The goal of the problem presentation was the stimulation of a student search or exploration promoted and guided by a problem which preceded the classroom discussion of certain topics (1968). Nietling found that:

> (1) The problem did stimulate student activity in the attempted solution of an unfamiliar problem; (2) the attempted solution of the problem did promote productive thinking, and (3) evidence that discovery of mathematical concepts had occurred was almost nonexistent.

His attempts to measure growth related to "the student's learning of the routine material of the course" revealed no significant differences.

From these studies, Goldstein concluded that:

> ...The data in most of the studies suggests that regularly assigned homework favors higher academic achievement, and a few of the best-designed experiments show this quite clearly.

Studies published during 1960 to 1977 tended not to use just a single test for the analysis. Rather, attitude measure as well as different levels of mathematics achievement measures were used. There were several homework studies that considered other than homework versus no homework related questions.

Maertens (1969) divided 342 third-grade pupils into three groups: one group had no homework; one had teacher-proposed homework, and one had researcher prepared

homework designed to give drill, verbal problems and problem situations requiring generalizations. Experimenter-prepared tests given after one year showed no difference between groups. No attitude differences were found (Maertens, 1968).

Maertens and Johnson (1972) divided 400 fourth, fifth and sixth grade students into three treatment groups. One group had no homework; one had homework with immediate feedback from parents, and one had homework with delayed feedback from parents. Experimenter-prepared tests were given after six weeks to assess computation and problem-solving abilities. Except for fifth grade problem-solving, both homework groups did significantly better than the no homework groups. On the fifth grade problem-solving test, the average favored the two homework groups, but the differences were not significant. The two homework groups did not differ on any test. No attitude differences were found. Note that both homework groups had some presented involvement to provide feedback on the problem solution.

Johnson, Maertens and Schooley (1969) made a short seven-day study on immediacy of feedback that did not involve homework. Using 68 fourth-grade pupils in three classes, grades on student-corrected papers indicated feedback immediately after each problem which was attempted was superior to feedback given after all problems were

attempted. The latter did not differ from giving feedback the day after all problems were attempted.

Gray and Allison (1971) used 64 sixth graders to compare no homework with drill-type homework. The homework group had parental involvement. After eight weeks no differences were found in computational skills. The authors concluded: "... The findings may indicate that drill-type homework is in fact unrelated to pupil growth in computational skills."

Several studies were made on homework grading. Small (1967) used 36 tenth-grade geometry students in a year-long study. One class had all homework assignments corrected while the other class had the assignment spot-checked. While the assignments were spot-checked, short daily tests were given. Although no significant differences between classes were found, 10 of the 13 class tests favored the class with assignments corrected daily. Austin and Austin (1974) made a seven-week study using 51 seventh and eighth-grade students. One group had every problem corrected and the other had a random half of the problems corrected. No differences were found between the two groups using class tests. The authors concluded that it may be a waste of teachers' time to grade every problem assignment.

Several studies attempted to adopt Page's (1958)

test grading study to homework grading. Using 2139 students in grade 7 through 12, Page randomly assigned students to one of the three groups in every class. Teachers wrote no comments on tests of students in one group, and they wrote standardized comments depending on the letter grade for the students in another group. Teachers wrote free comments on student tests in the other group. Teachers constructed tests during the school year uniformly--across grade and ability--favored the free comment groups over the no-comment group.

Austin (1976) made a similar study using 222 students in nine classes (grades 4, 9, and 10). In each class students were matched and randomly assigned to a comment or no-comment group. After six weeks the two groups were compared using teacher-prepared examinations. A significant difference favoring the comment group was found in two of nine classes, and when all students were pooled. Comments improved student achievement in geometry and one of two general mathematics classes and had no effect in algebra in one fourth-grade class.

Schoen and Kreye (1974) also made a study involving written comments on homework. For 12 weeks, several methods of writing comments were compared using 147 college-level prospective elementary education majors. No differences were found on test scores. This may not

generalize to pre-college students as large class sizes and college students were used.

The following conclusions--grouped by pre- and post-1950 empirical support, can be drawn from the available studies.

Conclusion based on pre-1950 studies:

1) The home environment is an important factor in the effectiveness of homework (Brook, 1916).

2) Required homework may be preferable to voluntary homework (Di Napoli, 1937).

3) No homework in one grade can adversely affect performance in subsequent grades (Crawford, 1937).

Conclusion primarily based on post-1950 studies:

1) Homework seems preferable to no-homework at least for grades 4 through 10.

2) The effect of homework may be cumulative in nature (Hines, 1957).

3) Parental involvement alone does not ensure the effectiveness of homework. Maertens and Johnson (1972) showed significant differences favoring homework group with parental involvement. Gray and Anderson (1971) found no difference. See (1).

4) Homework seems generally to improve computational skills. Kock (1965) and Maertens (1972) found significant differences favoring the homework group. Maertens (1969) found no difference.

5) The effect of homework on problem-solving skills is less clear. Maertens and Johnson (1972) found significant differences favoring the homework group. Kock (1965) and Maertens (1969) found no differences.

6) There is little information on the length of homework assignments. Kock (1965) found longer assignments improve computation skills but did not seem to affect problem-solving.

7) No study had found a relationship between homework and attitude toward mathematics (Maertens, 1968).

8) Every homework problem need not be graded (Austin and Austin, 1974).

9) Routine drill homework may not be of much value (Gray, 1971).

CHAPTER III

THE STUDY

This chapter presents the methods and design of the study in the following main sections:

1. The purpose of the study
2. The hypotheses
3. The experimental procedures
4. Statistical analysis
5. Summary

THE PURPOSE OF THE STUDY

The purpose of the study was to address the issue of whether teachers' written comments on homework papers are useful and worth the effort. To help respond to this issue, some students received teacher comments; others received no comments, and an independent test measured whether these groups thereafter behaved differently from one another. It was also the aim of this study to examine the relationship between the comments effect and the students' grade levels.

THE MAJOR HYPOTHESES

The major hypotheses are as follows:

Treatment

H_1: There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average gain on teacher made tests.

Grade

H_2: There is no significant difference in the mean grade levels in fourth, sixth and ninth grade mathematics students between the experimental group and the control group as measured by the average gain on teacher made tests.

Interaction

H_3: There is no significant interaction between the treatment and grade levels of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average score on teacher made tests.

THE MINOR HYPOTHESES

In order to test the treatment effect on students' ability level the investigator obtained the previous term mathematics score (out of 100 points) for each of the involved students. Three minor hypotheses for each

main hypothesis were tested. The students were divided into three groups to test for which group the treatment worked best. These groups were:

Group 1. The students who scored up to 55 points on the previous term mathematics examination.

Group 2. The students who scored between 55 and 75 points on the previous term mathematics examination.

Group 3. The students who scored more than 75 points on the previous term mathematics examination.

The minor hypotheses are as follows:

Treatment

$H_{1.1}$: There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students (Group 1) in the experimental group and the control group as measured by the average score on teacher made tests.

$H_{1.2}$: There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students (Group 2) in the experimental group and the control group as measured by the average score on teacher made tests.

$H_{1.3}$: There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students (Group 3) in the experimental group and the control group as measured by the average score on teacher made tests.

Grade

$H_{2.1}$: There is no significant difference in the mean grade levels of fourth, sixth and ninth grade mathematics students (Group 1) in the experimental group and the control group as measured by the average score on teacher made tests.

$H_{2.2}$: There is no significant difference in the mean grade levels of fourth, sixth and ninth grade mathematics students (Group 2) in the experimental group and the control group as measured by the average score on teacher made tests.

$H_{2.3}$: There is no significant difference in the mean grade levels of fourth, sixth and ninth grade mathematics students (Group 3) in the experimental group and the control group as measured by the average score on teacher made tests.

Interaction

$H_{3.1}$: There is no significant interaction between the treatment and the grade levels of fourth, sixth

and ninth grade mathematics students (Group 1) in the experimental group and the control group as measured by the average score on teacher made tests.

$H_{3.2}$: There is no significant interaction between the treatment and the grade levels of fourth, sixth and ninth grade mathematics students (Group 2) in the experimental group and the control group as measured by the average score on teacher made tests.

$H_{3.3}$: There is no significant interaction between the treatment and the grade levels of fourth, sixth and ninth grade mathematics students (Group 3) in the experimental group and the control group as measured by the average score on teacher made tests.

THE EXPERIMENTAL PROCEDURE

Permission was received from the Saudi Arabian Ministry of Education to conduct the study in the city of Mecca, during the second semester, 1981.

The investigator secured directories for all fourth, sixth and ninth grade mathematics teachers. Using a table of random numbers, and coding the directories concerned, six teachers in each grade were selected and information recorded about each teacher so selected.

To secure cooperation of the teachers selected for

the study, the investigator talked to each teacher, in his school, during one of his free periods. During each meeting the investigator described the study and the teacher's role. Some teachers were not personally acquainted with the experimenter and had been carefully informed that their contribution was to be entirely voluntary and unrewarded. Their cooperative spirit made recruitment far more pleasant than might be imagined.

Two classes for each teacher were randomly assigned, one into the experimental group and the other into the control group. Since two classes for each of the selected teachers were randomly selected from the eligible classes, the classes are themselves a random sample of the fourth, sixth and ninth grade mathematics classes in the city of Mecca.

The teachers administered the pre-tests to all the students. The teachers in each grade participated in preparing this test. They reported each student's score to the investigator.

The treatment was then administered. The No-Comment group received no marks beyond those for grading. The homework papers in this group were returned with no information beyond the numerical score. Homework papers in the Experimental group had teacher's comments, whatever comments the teacher felt were appropriate.

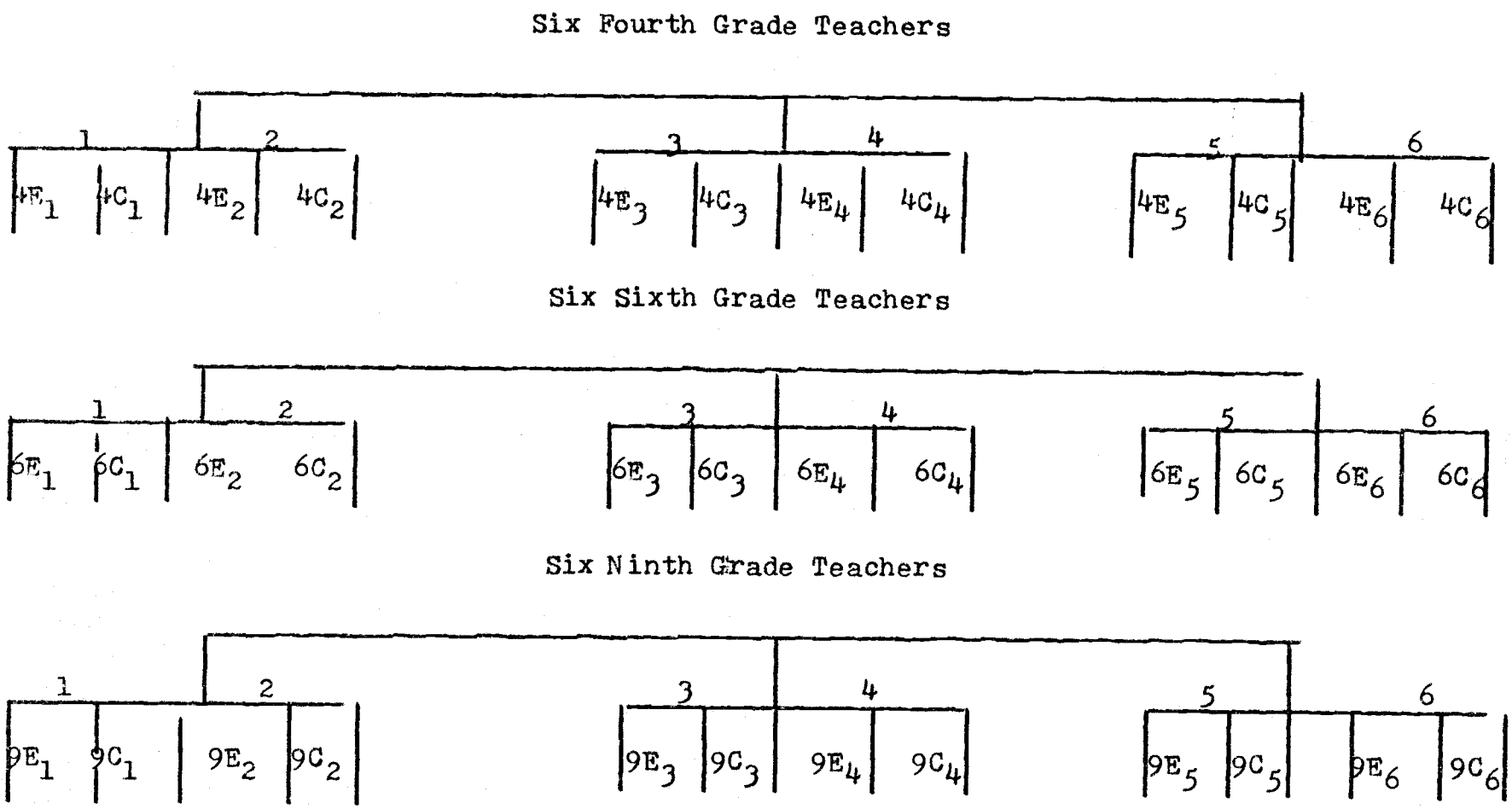

Figure 1. Sample Selection Procedure

If a student missed a problem, the teacher marked where the error occurred, and either gave or indicated how to arrive at a correct solution. If a student had a clever solution, the teacher acknowledged this.

The teachers were instructed:

> Forget you are going to report the comment to anyone else. Write anything that occurs to you in the circumstances. Please remember only one thing, the reason we teachers comment on homework papers is that we believe it will in some way encourage the student to perform well in the future.

An experimental assumption was that teachers' comments affect students' performance. The criterion of treatment effect was therefore the post-test administered in each classroom. This test was also teacher-made. It was a different form of the pre-test, the teachers reported each student score on both tests, and returned the report to the investigator.

Population

The population for the study consisted of twelve classes in each of three grade levels: six classes assigned as the experimental group and the other six classes assigned as the control group. The number of students in each grade is described in the following table. Tables that display the sample in detail are found in Appendix B.

Table 1. Number of Students in Each Grade Level

Grade	Fourth	Sixth	Ninth	Total
Experimental Group	184	165	178	527
Control Group	189	162	182	533
Total	373	327	360	1,060

The Population in Group 1

The number of students who scored less than 55 points in the previous term mathematics examination was 541. Table 2 describes the number in each grade level. Tables that display the population in detail are found in Appendix B.

Table 2. Summary of the Population in Group 1

Grade	Fourth	Sixth	Ninth	Total
Experimental Group	82	90	94	266
Control Group	94	81	100	275
Total	176	171	194	541

The Population in Group 2

The number of students who scored between 55 and 75 points in the previous term examination was 374. Table 3 describes the number of the students in each grade level.

Tables that display the population in detail are found in Appendix B.

Table 3. Summary of the Population in Group 2

Grade	Fourth	Sixth	Ninth	Total
Experimental Group	59	54	71	184
Control Group	59	63	68	190
Total	118	117	139	374

The Population in Group 3

The total number of the students who scored more than 75 points in the previous term examination was 145. Table 4 describes the number in each grade level. Tables that display, in detail, the distribution of these students, are found in Appendix B.

Table 4. Summary of the Population in Group 3

Grade	Fourth	Sixth	Ninth	Total
Experimental Group	43	21	13	77
Control Group	36	18	14	68
Total	79	39	27	145

Classroom Teachers

As already mentioned, there were 12 classes in each

grade level, and each teacher taught two classes; one as experimental group and the other one as control group, so there was a total of eighteen teachers involved. There were six teachers in each grade level. A brief information resume about each one was recorded in the directories which the investigator had obtained from the Educational Bureau.

Fourth Grade Teachers

Teacher 1 is 32 years old. He has a high school diploma and enrolled for two years in the Teachers Institute. Since he received his diploma he has taught fourth and fifth grade mathematics for ten years.

Teacher 2 is 30 years old. He earned the same degree as teacher 1 and has taught sixth grade mathematics for eight years.

Teacher 3 is 30 years old. He graduated from the previous system of the Teachers Institute. In that system, a student who finished 9th grade can enroll in the institute for four years and become a teacher. He has taught fourth and fifth grade mathematics for ten years.

Teacher 4 is 31 years old. He graduated from the same institute as teacher 3. He has taught mathematics for nine years. He has taught third grade mathematics for three years.

Teacher 5 is 30 years old. He has taught fourth

grade mathematics for ten years. He also taught fourth grade general science. He had his high school diploma and two years enrollment in the Teachers Institute.

Teacher 6 is 32 years old. He studied in the Teachers' Institute after he received his high school diploma. He has taught fourth grade mathematics for 3 years.

Sixth Grade Teachers

Teacher 1 is 35 years old. He has taught fourth grade mathematics for 5 years. He also teaches sixth grade mathematics classes.

Teacher 2 is 33 years old. He was a third grade teacher in a self-contained classroom. He also has taught sixth grade mathematics for 4 years.

Teacher 3 is 29 years old. He also has taught fourth grade mathematics for 5 years. He teaches general science in two fifth grades and mathematics in three sixth grade classes.

Teacher 4 is 30 years old. He also has taught third grade classes for 3 years. He has taught fourth and fifth grade mathematics. He has taught sixth grade mathematics for two years.

Teacher 5 is 32 years old. He has been a fourth grade mathematics teacher for 5 years. He has taught

sixth grade mathematics for 3 years.

Teacher 6 is 36 years old. He has been a sixth grade mathematics teacher for 3 years.

Ninth Grade Teachers

Teacher 1 is 34 years old. He had a high school diploma and became a fourth grade mathematics teacher after two years enrollment in the Teachers' Institute. He earned his B.A. four years ago. Ever since he has been a ninth grade mathematics teacher.

Teacher 2 is 32 years old. He is from Egypt. He earned his B.A. degree from Cairo University. He came to Saudi Arabia in 1978. He has taught ninth grade mathematics for three years. He expects to spend two more years teaching tenth grade mathematics.

Teacher 3 is 28 years old. He earned a B.A. degree from King Abdul-Aziz University in mathematics. He has taught seventh and eighth grade mathematics for three years. He has been teaching ninth grade mathematics for two years.

Teacher 4 is 32 years old. He is from Jordan. He earned a B.A. degree with a major in mathematics and a minor in science education. He came to Saudi Arabia in 1978. Since then he has been a ninth grade mathematics teacher.

Teacher 5 is 29 years old. He earned his B.S.

with a major in physics and mathematics and a minor in education. He has taught ninth grade mathematics, and eighth grade general science.

Teacher 6 is 32 years old. He earned his B.S. from Cairo University with a major in accounting and a minor in mathematics. He has been a ninth grade mathematics teacher for two years. He expects to teach accounting in a commercial high school for the next three years.

RESEARCH DESIGN

Classical experimental design allows the variation of a single variable at a time, holding all other conditions constant. Complex behaviors, however, usually will not be fitted into such an artificial situation. Furthermore, factors influencing behaviors frequently interact to produce differences that do not occur when only one factor is free to vary at a time. For example, there might be an interaction between intelligence and method of teaching, such that authoritarian methods work better with low-intelligence subjects and democratic methods work better with high intelligence subjects. A study of the two teaching methods which holds intelligence constant will lead to a misleading conclusion, just as would a study allowing intelligence to vary but holding the teaching methods constant. To permit studies where more than one

factor is free to vary at a time, Split-Plot designs have become increasingly popular. The term Split-Plot comes from agricultural experimentation in which a single level of one treatment is applied to a relatively large plot of ground (the whole plot), but all levels of a second treatment are applied to subplots within the whole plot.

There are many research problems in the behavioral science where Split-Plot designs are especially appropriate. A general problem inherent in all behavioral research is subject heterogeneity. Differences among subjects are often such as to obscure treatment effects. A Split-Plot design offers the advantage of controlling subject heterogeneity. These designs have several advantages over the classical experimental design:

1. They permit the testing of several hypotheses simultaneously, rather than having to conduct a series of single variable experiments to study the effects of different variables on, for example, learning.

2. They permit the conduct of only one experiment to answer several complex questions at once, such as: What effect does X_1-type teacher have on learning achievement when using X_2-type methods in X_3-length classes?

3. Where interaction between two or more variables simultaneously makes a difference, it reveals this difference.

It should be clear then, that this design yields far more information than could be gathered from a single-variable experiment with the available subjects, say,one hundred, but concerned with either variable A or variable B alone, it yields even more information than could be obtained from two such experiments together using twice as many subjects. The Split-Plot design and the method of analysis appropriate to it, which are due to R. A. Fisher, have often been described as among the most important contribution to experimental technique in recent decades.

When studying the effect of teacher comments and achievement, not only is the effect itself of interest, but the effect of these comments on achievement at a particular grade level is of interest as well. This design allows us not only to assess the effect of teacher comments on student's achievement, but also allows us to draw some conclusions as to the relationship between teacher comments and student's grade level. These are really two independent variables. The first one (experimental versus control), and the second variable is the grade level. Table 5 represents the design matrix.

Table 5. Design-Matrix

	Fourth Grade	Sixth Grade	Ninth Grade
Experimental Group			
Control Group			

Data Collection

All the students were pretested with the achievement examination. Each teacher selected the appropriate time, any day in the first week. The achievement posttest was administered to all the students on the eleventh week. The teachers reported each student's score on those tests.

In summary, the raw data consisted of 1,060 observations for measuring mean achievement difference. Detailed descriptions of the students' distribution in each grade level are in Appendix B. The data were summarized and cardpunched in preparation for statistical analysis on the computer at Oregon State University. Analysis of variance was used to analyze the data.

STATISTICAL ANALYSIS OF THE DATA

The basic statistical technique used to test the null hypotheses in this design is the analysis of variance, Table 6. The calculation utilizes the analysis procedures result in an F-ratio. This ratio is the ratio of the variable between groups to the variable within groups.

Since there are two independent variables in this study, there must be three F-ratios, one for each main variable and one for the interaction. To compute an F-ratio, these two variables must be estimated. In the process of obtaining an estimate of these two variables, a statistical term called the sum of squares SS for each source of variance must be calculated. From the SS, calculate the mean square MS by dividing the sum of the squares by the appropriate degree of freedom, df. The mean square is an estimate of the variance. The value of the F-ratio for each source is obtained by dividing each mean square by the mean square error.

Hypotheses Testing

To test H_2: there is no significant grade effect, use

$$F_{\alpha,\ c,\ e} = \frac{MSG}{MST(G)}$$

To test H_3: there is no significant interaction between treatment and grade, use

$$F_{\alpha,\ a,\ b} = \frac{MS(Tr \times G)}{MSE}$$

If this is not significant, i.e., they do not interact, then test for treatment effect by

$$F_{\alpha,\ d,\ e} = \frac{MSTr}{MSE}$$

A graphic plot will be necessary to show where the interaction occurred.

Table 6. Analysis of Variance

Source of Variation	dF	Sum of Square	Mean Square	F
Grade	a	G	$\frac{G}{a}$	$\frac{\text{MS grade}}{\text{MS teacher/grade}}$
Teacher/grade	b	H	$\frac{H}{b}$	- - -
Treatment	c	I	$\frac{I}{c}$	$\frac{\text{MS treatment}}{\text{MS error}}$
Treatment x grade	d	J	$\frac{J}{d}$	$\frac{\text{MS interaction}}{\text{MS error}}$
Error	e	K	$\frac{K}{e}$	- - -
Total	f	M	$\frac{M}{f}$	- - -

SUMMARY

This study was conducted to determine if teacher comments on student homework papers significantly increased student achievement in mathematics. The nested Split-Plot design was chosen for the experimental design.

Two forms of the course achievement examination measured the students' achievements. Thirty-six mathematics classes from three different grades, 12 classes for each fourth, sixth and ninth grades, were randomly assigned to the two groups. The homework papers in the control group were graded by simply marking problems right or wrong. Homework papers in the experimental group had teacher comments on students' homework. For ten weeks the teachers collected homework papers three times weekly. Both the control and the experimental group were pre- and post-tested. The statistical technique used to test the hypotheses in this design was the analysis of variance. This chapter presented detailed discussion of the methods and procedures used in the study.

CHAPTER IV

RESULTS OF THE STUDY

This chapter presents the findings of the study. The first section deals with the findings related to the major hypotheses; the second section deals with the findings related to the minor hypotheses.

FINDINGS RELATED TO THE MAJOR HYPOTHESES

Hypothesis One (Treatment)

There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average gain on teacher made tests.

To test this hypothesis a two-way analysis of variance was used. Table 7 presents the results of the analysis. As shown in the table, the F-value was significant at the 0.05 level.

Since the calculated F-value was greater than the tabular F-value, the hypothesis was rejected. From this it was concluded that there was a significant difference in the mean achievement between the experimental group and the control group as measured by the average gain on teacher made tests. Table 8 below shows that there was an improvement in the mean achievement in favor of the

Table 7. Analysis of Variance for the Major Hypotheses

Source of Variation	DF	Sum of Square	Mean Square	F
Grade	2	71.121	35.56	0.1366
Teacher/Grade	15	3904.700	260.31	-
Treatment	1	28695.86	28695.86	272.3*
Treatment x Grade	2	231.06	115.53	1.096
Error	1039	109478.68	105.37	-
Total	1059	142381.43	-	-

*Significant at 0.05 level.

Table 8. Mean Gain in Achievement by Grade

Grade	Fourth	Sixth	Ninth	Average
Control Group	52.10	53.389	52.72	52.7
Experimental Group	63.85	63.39	62.33	63.2
Average	57.8	58.38	57.5	58.0

experimental group. This progress in the students' achievement is attributed to the treatment effect. Since the students in the fourth, sixth and ninth grade who received teachers' comments (experimental group) scored higher than the students in the fourth, sixth and ninth grade who received no comments (control group), hypothesis one was rejected.

Table 8 shows that there is no significant grade effect. The means in the fourth, sixth and ninth grade were about the same.

Hypothesis Two (Grade)

There is no significant difference existing in the mean grade level of the fourth, sixth and ninth grade mathematics students between the experimental group and the control group as measured by the average gain on teacher made tests.

The analysis of this hypothesis was similar to hypothesis one. Table 7 shows that the F-value of 0.1366 was not significant at the 0.05 level.

Since the calculated F-value was less than the tabular F-value, hypothesis two was not rejected. From this it was concluded that there was no significant grade level effect.

Table 8 shows that fourth, sixth and ninth grade students improve their scores on the average of 58 points. The grade level effect was not significant. This indicated that hypothesis two was not rejected. It was concluded that there is no significant difference existing in the mean grade level of fourth, sixth and ninth grade mathematics students between the experimental group and the control group as measured by the average gain on teacher made tests.

Hypothesis Three (Interaction)

Hypothesis three stated that there is no significant interaction between the treatment effect and grade levels.

To test this hypothesis a two-way analysis of variance was used. Table 7 presents the results of the analysis. As shown in the table, the interaction F-value was not significant at the 0.05 level.

The analysis of this hypothesis is similar to the analysis of hypothesis one. Table 7 shows that the F-value of 1.096 was not significant at the 0.05 level.

Since the calculated F-value (1.096) was less than the tabular F-value (2.99), the hypothesis was not rejected. There was no significant interaction between the treatment and grade levels found to exist.

Consider the data in Table 8. First consider the treatment effect, the students improve their scores on the average of 63.2 points under the written comments method and improve their scores on the average of 52.7 points under the No-Comments Method. Thus the way teachers correct homework was significant. Second consider the grade level effect. Since the students in the fourth, sixth and ninth grade improve their score on the average of 58 points, the grade level effect was not significant. Schematically this is represented in Fig. 2. Not only the comment method is more effective than the no-comment

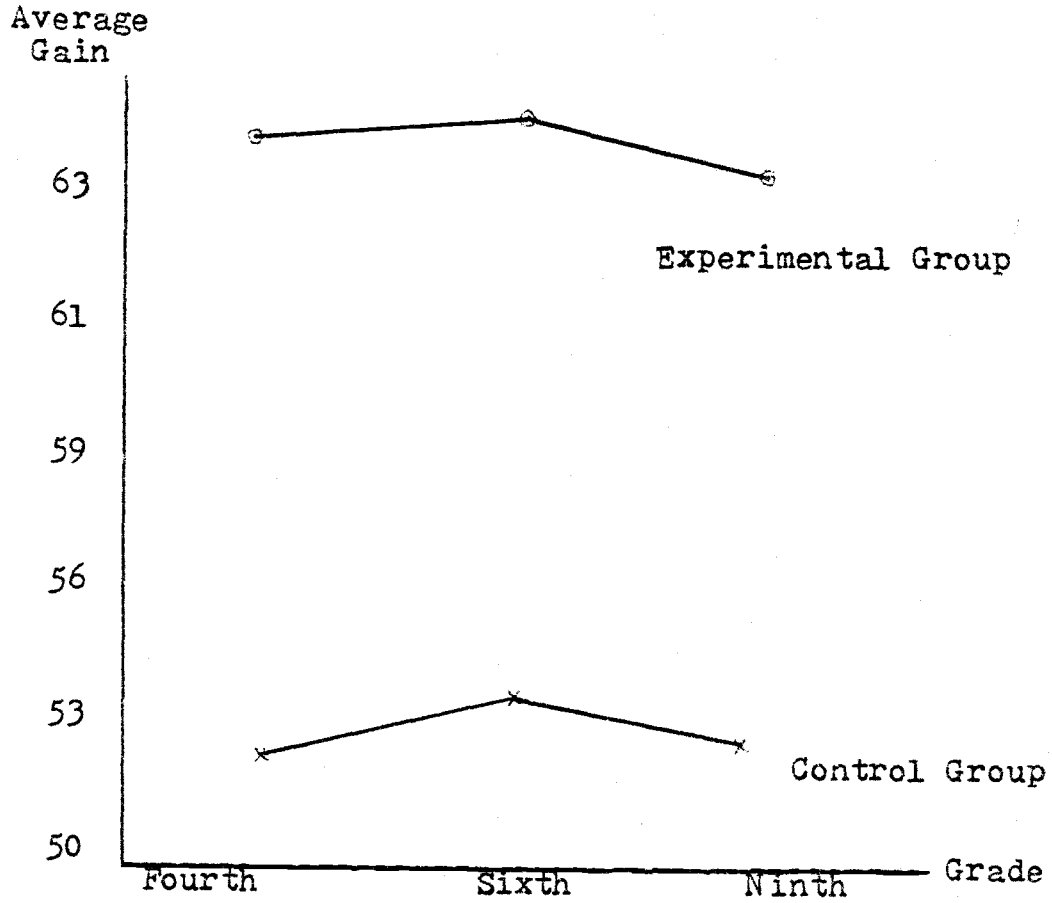

Figure 2. Interaction Between the Treatment Effect and the Grade

method, but it was equally effective in each grade level.

With data as complex as these, it is a good idea to plot them as in Fig. 2. Although they go up and down, they do so together; that is, they are roughly parallel. Therefore, there is no significant interaction, between the treatment effect and the grade levels.

FINDINGS RELATED TO THE MINOR HYPOTHESES

Three minor hypotheses, for each major one, were tested. These hypotheses dealt with the students who scored up to 55 points, between 55 and 75 points, and the students who scored more than 75 points in the previous term mathematics examination.

The investigator secured the score for each student in the previous term mathematics examination. The students were then divided into three groups according to their score in the previous term mathematics examination.

Group 1. Students who scored up to 55 points.

Group 2. Students who scored between 55 and 75 points.

Group 3. Students who scored more than 75 points.

GROUP 1

Treatment

$H_{1.1}$: There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average gain on teacher made tests.

To test this hypothesis a two-way analysis of variance was used. Table 9 presents the results of the analysis. As shown in the table, the F-value was

Table 9. Analysis of Variance for the Minor Hypothesis in Group 1

Source of Variance	DF	Sum of Square	Mean Square	F
Grade	2	160.962	80.481	1.386
Teacher/Grade	15	870.898	58.060	-
Treatment	1	17433.750	13433.75	196.05*
Treatment x Grade		240.898	120.449	1.758
Error	520	35641.116	68.52	-
Total	540	50337.624	-	-

*Significant at 0.05 level.

significant at the 0.05 level.

Since the calculated F-value was greater than the tabular F-value, the hypothesis was rejected. From this it was concluded that there was a significant difference in the mean achievement between the experimental group and the control group as measured by the average gain on teacher made tests for Group 1.

Table 10 showed an improvement in the mean in favor of the experimental group. This progress in the students' achievement is attributed to the comment method.

The students in the fourth, sixth and ninth grade, who received teachers' comments on homework papers (experimental group 1) scored higher than the students in the fourth, sixth and ninth grade, who received no

comments (control group 1), hypothesis one was rejected.

Grade

$H_{2.1}$: There is no significant difference existing in the mean grade levels of the fourth, sixth and ninth grade mathematics students, between the experimental group and the control group as measured by the average gain on teacher made tests for group 1.

To test the hypothesis a two-way analysis of variance was used. Table 9 presents the results of the analysis. As shown in the table, the F-value was not significant at the 0.05 level.

Since the calculated F-value was less than the tabular F-value, the hypothesis was not rejected. From this it was concluded that there was no significant grade level effect. The means in all grades were about the same.

Table 10 shows that fourth, sixth and ninth grade students improve their scores on the average of 51 points. The main effect of grade levels was not significant for group 1. This indicated that hypothesis two was not rejected. There is no significant difference existing in the mean grade levels of fourth, sixth and ninth grade mathematics students, between the experimental group and the control group as measured by the average gain on the teacher made tests.

Table 10. Mean Average Gain in Achievement by Grade for Group 1

Grade	Fourth	Sixth	Ninth	Average
Control Group	44.97	46.74	46.10	45.94
Experimental Group	56.65	57.11	54.62	56.15
Average	50.41	52.19	50.26	51

Interaction

$H_{3.1}$: There is no significant interaction between the treatment and grade levels for Group 1.

To test the hypothesis a two-way analysis of variance was used. Table 9 represents the results of the analysis. As shown in the table, the interaction F-value was not significant at the 0.05 level.

Since the calculated F-value was less than the tabular F-value, the hypothesis was not rejected. There was no significant interaction between the treatment and the grade levels found to exist.

Consider the data in Table 10. First consider the treatment effect, the students improve their score on the average of 56.15 points under the written comments method and on the average of 45.94 points under the no-comments method. Thus the way teachers correct homework was significant. Second consider the grade level effect.

Since the students in the fourth, sixth and ninth grade improve their score on the average of 51 points, the grade level effect was not significant. Schematically this is represented in Fig. 3. Not only is the comments method more effective than the no-comment method, but it is equally effective in each grade level.

Fig. 3 shows that the students in the experimental group gained more than the students in the control group. They did so in all grade levels. This indicated that there is no significant interaction between the treatment effect and grade levels.

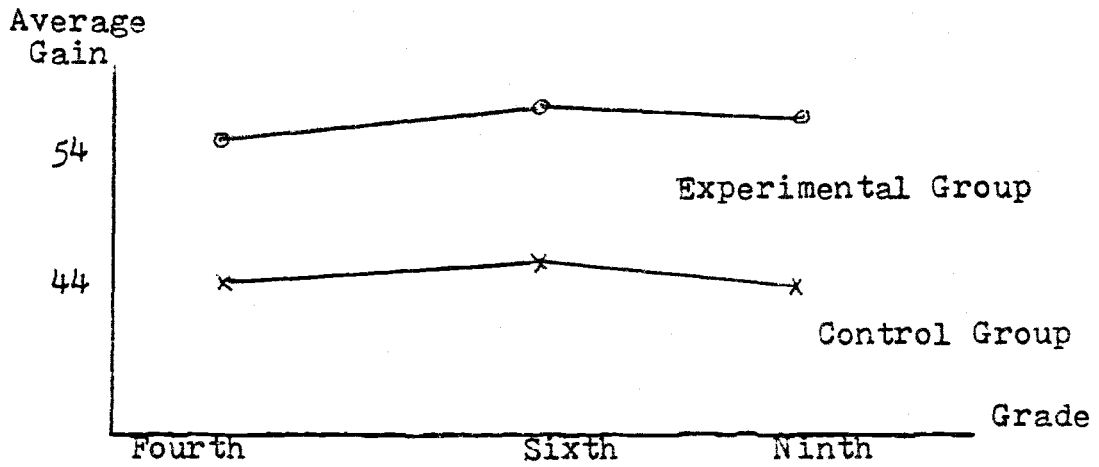

Figure 3. Interaction Between the Treatment Effect and the Grade for Group 1

GROUP 2

In this section three minor hypotheses dealt with the students who scored between 55 and 75 points in the

previous term mathematics examination and will be discussed.

Treatment

$H_{1.2}$: There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average gain on teacher made tests.

To test this hypothesis a two-way analysis of variance was used. Table 11 presents the results of the analysis. As shown in the table, the F-value was significant at the 0.05 level.

Since the calculated F-value was greater than the tabular F-value, the hypothesis was rejected. From this it was concluded that there was a significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average gain on teacher made tests for Group 2.

Table 12 shows improvement in the mean in favor of the experimental group. This progress in the students' achievement is attributed to the effects of the comment method.

The students in the fourth, sixth and ninth grade

who received teachers' comments on mathematics homework papers (experimental group) improve more than the students who received no-comments on mathematics homework papers (control group).

Table 12 shows that there was no significant grade level effect. The means for the fourth, sixth and ninth grades were about the same.

Table 11. Analysis of Variance for Minor Hypothesis in Group 2

Source of Variation	DF	Sum of Squares	Mean Square	F
Grade	2	443.601	221.8	2.478
Teacher/grade	15	1342.85	89.5	-
Treatment	1	7313.17	7313.17	256.55*
Treatment x grade	2	42.52	21.17	0.7397
Error	353	10145.611	21.26	-
Total	373	19347.778	-	-

*Significant at 0.05 level.

Table 12. Mean Average Gain in Achievement by Grade for Group 2

Grade	Fourth	Sixth	Ninth	Average
Control Group	57.37	58.46	59.41	58.41
Experimental Group	66.17	68.61	68.94	67.91
Average	61.77	63.535	64.17	63

Grade

$H_{2.2}$: There is no significant difference existing in the mean grade levels of the fourth, sixth and ninth grade mathematics students, between the experimental group and the control group as measured by the average gain on the teacher made tests.

To test the hypothesis a two-way analysis of variance was used. Table 11 presents the results of the analysis. As shown in the table, the F-value was not significant at the 0.05 level.

Since the calculated F-value was less than the tabular F-value, the hypothesis was not rejected. From this it was concluded that there was no significant grade level effect.

Table 12 shows that fourth, sixth and ninth grade students improve their scores on the average of 63 points. The main effect of grade levels was not significant. The means for all grade levels were about the same. There is no significant difference existing in the mean grade levels of fourth, sixth and ninth grade mathematics students, (Group 2), between the experimental group and the control group as measured by the average gain on the teacher made tests.

Interaction

$H_{3.2}$: There is no significant interaction between the

treatment and the grade level for group 2.

To test the hypothesis a two-way analysis of variance was used. Table 11 presents the results of the analysis. As shown in the table, the interaction F-value was not significant at the 0.05 level.

Since the calculated F-value was less than the tabular F-value, the hypothesis was not rejected. There was no significant interaction between the treatment and the grade levels found to exist.

Considering the data in Table 12, first consider the treatment effect. The students improve their scores on the average of 67.91 points under the written comments method and on the average of 58.41 points under the no-comment method. Thus the way teachers correct homework papers was significant.

Second consider the grade level effect. Since the students on the fourth, sixth and ninth grade improve their scores on the average of 63 points, the grade level effect was not significant. All means in the fourth, sixth and ninth grades are about the same. Schematically this is represented in Fig. 4. The figure shows that not only the comment method is more effective than the no-comment method, but it is equally effective in each grade level. From this it was concluded that although the students in the experimental group gained more than the students in

the control group, they did so in all grade levels for group 2. This indicated that there is no significant interaction between the treatment and grade levels.

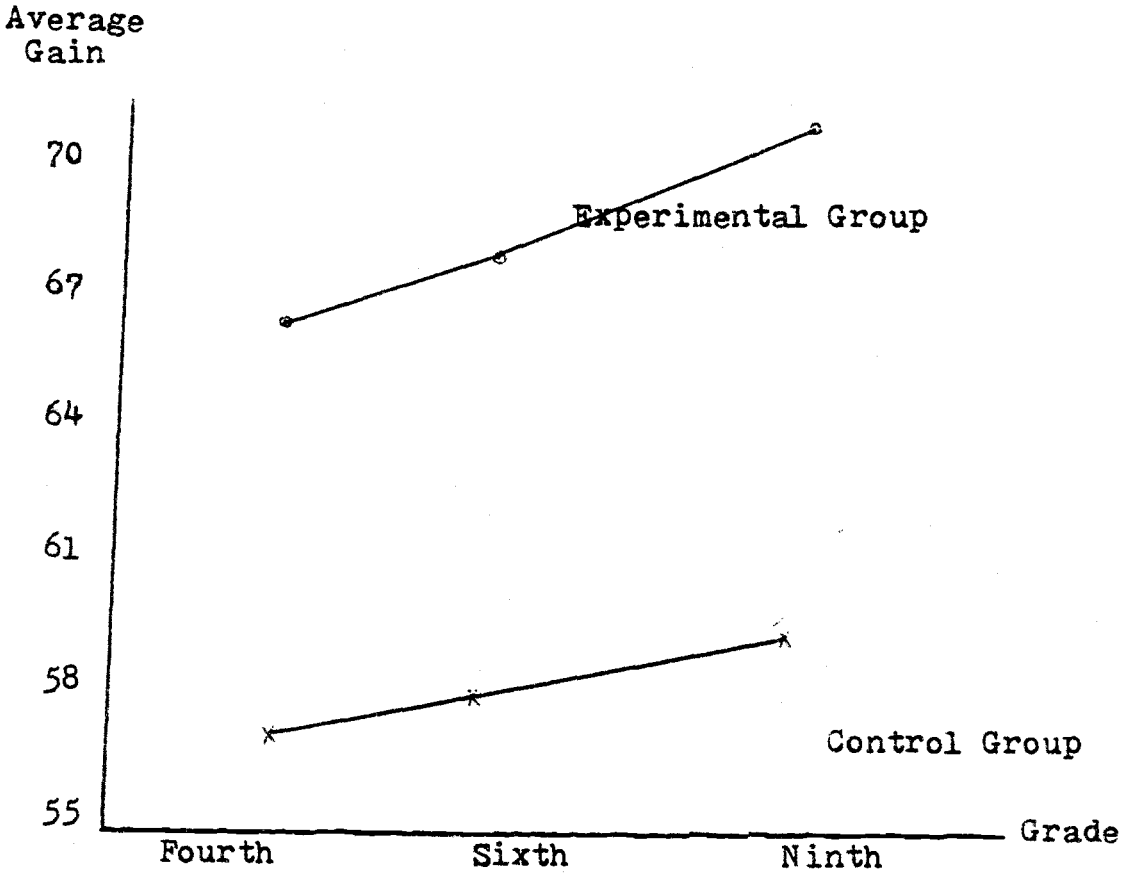

Figure 4. Interaction Between the Treatment Effect and the Grade for Group 2

GROUP 3

In this section three hypotheses deal with the students who scored more than 75 points in the last term examination.

Treatment

$H_{1.3}$: There is no significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students in the experimental group for group 3 and the control group as measured by the average gain on teacher made tests.

To test this hypothesis a two-way analysis of variance was used. Table 13 represents the results of the analysis. As shown in the table, the F-value was significant at the 0.05 level.

Since the calculated F-value was greater than the tabulated F-value, the hypothesis was rejected. From this it was concluded that there was a significant difference in the mean achievement between the experimental group and the control group for group 3.

Table 14 shows improvement in the mean in favor of the experimental group. This progress in the students' achievement is attributed to the effect of the comments method.

The students in the fourth, sixth and ninth grade who receive teachers' comments on mathematics homework papers (experimental group) improved more than the students in the fourth, sixth and ninth grade who received no-comments on mathematics homework papers (control group). This led to rejection of the hypothesis.

Table 13. Analysis of Variance for the Minor Hypotheses in Group 3

Source of Variation	DF	Sum of Square	Mean Square	F
Grade	2	325.874	162.937	3.23
Teacher/Grade	14	705.903	50.42	-
Treatment	1	4014.818	4014.818	111.06*
Treatment x grade	2	0.840	0.420	0.0115
Error	125	4518.657	36.149	-
Total	144	4566.092	-	-

*Significant at 0.05 level.

Table 14. Mean Average Gain in Achievement by Grade for Group 3

Grade	Fourth	Sixth	Ninth	Average
Control Group	62.08	65.55	67.50	65.04
Experimental Group	74.42	76.90	81.54	77.62
Average	68.25	71.22	74.52	71

Grade

$H_{2.3}$: There is no significant difference existing in the mean grade levels of the fourth, sixth and ninth grade mathematics students (Group 3), between the experimental group and the control group as measured by the average gain on teacher made tests.

To test the hypothesis a two-way analysis of

variance was used. Table 13 presents the results of the analysis. As shown in the table, the F-value was not significant at the 0.05 level.

Since the calculated F-value was less than the tabular F-value, the hypothesis was not rejected. From this it was concluded that there was no significant grade level effect. Table 14 shows that fourth, sixth and ninth grade students improve their scores on the average of 71 points. The main effect of grade level was not significant for group 3. The means for all grade levels were about the same. This indicated that the hypothesis was not rejected. There is no significant difference existing in the mean grade levels of fourth, sixth and ninth grade mathematics students, between the experimental group and the control group as measured by the average gain on the teacher made tests.

Interaction

$H_{3.3}$: There is no significant interaction between the treatment and the grade levels of the fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average gain on teacher made tests for group 3.

To test the hypothesis a two-way analysis of variance was used. Table 13 presents the results of the analysis. As shown in the table, the F-value was not

significant at the 0.05 level.

Since the calculated F-value was less than the tabular F-value, the hypothesis was not rejected. There was no significant interaction between the treatment and the grade levels found to exist.

Consider the data in Table 14. First consider the treatment effect. The students improve their scores on the average of 77.62 points under the written comment method and on the average of 65.04 points under the no-comment method. Thus the way teachers correct homework papers was significant.

Second consider the grade level effect. Since the students in the fourth, sixth and ninth grade improve their scores on the average of 71 points, the grade level effect was not significant. All means in the fourth, sixth and ninth grade are about the same. Schematically this is represented in Fig. 5. The figure shows that not only the comments method is more effective than the no-comment method, but it is equally effective in each grade level. From this it was concluded that although the students in the experimental group gained more than the students in the control group, they did so in all grade levels for group 3. This indicated that there is no significant interaction between the treatment and the grade levels.

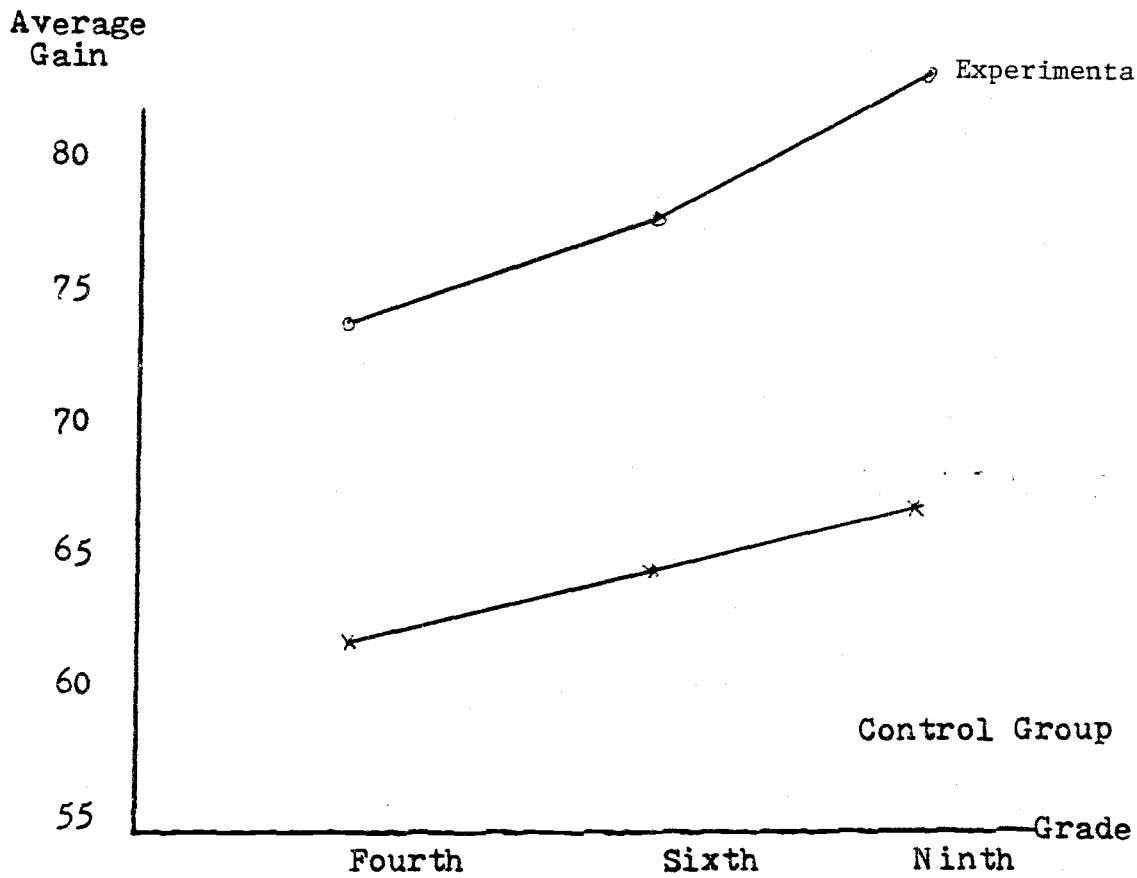

Figure 5. Interaction Between the Treatment Effect and the Grade for Group 3

SUMMARY OF THE RESULTS

The Major Findings

The findings for each of the major hypothesis are as follows:

1. There is a significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average score on teacher made tests.

2. There is no significant difference in the mean

grade levels of fourth, sixth and ninth grade mathematics students in the experimental group and control group as measured by the average score on teacher made tests.

3. There is no significant interaction between the treatment and the grade levels of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average score on teacher made tests.

The Minor Findings

Group 1.

1. There is a significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average score on teacher made tests.

2. There is no significant difference in the mean grade levels of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average score on teacher made tests.

3. There is no significant interaction between the treatment and the grade level of the fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average gain on teacher made tests.

Group 2.

1. There is a significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average score on teacher made tests.

2. There is no significant difference in the mean grade level of fourth, sixth and ninth grade mathematics students in the experimental and control groups as measured by the average score on teacher made tests.

3. There is no significant interaction between the treatment and grade levels of the fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average score on teacher made tests.

Group 3.

1. There is a significant difference in the mean achievement of fourth, sixth and ninth grade mathematics students in the experimental group and the control group as measured by the average score on teacher made tests.

2. There is no significant difference in the mean grade level of fourth, sixth and ninth grade mathematics students in the experimental and the control groups as measured by the average score on teacher made tests.

3. There is no significant interaction between the treatment and grade levels of the fourth, sixth and ninth

grade mathematics students in the experimental group and the control group as measured by the average score on teacher-made tests.

The results of the findings related to the minor hypotheses imply that there is no difference in the effect of the treatment among ability levels; that is, that each of the three ability levels performed similar to the group as a whole, and thus was effective at all ability levels.

CHAPTER V

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

This chapter is divided into three main sections. The first section reviews the entire study. The second section presents the conclusions and relates the findings to research. The third section presents recommendations for further study and recommendations for educational practice.

SUMMARY OF THE STUDY

The Problem

Each year mathematics teachers spend many hours writing comments on homework papers being returned to students apparently in the belief that their words will produce some results in students' performance superior to that obtained without such words. Yet on this point solid experimental evidence, obtained under classroom conditions, has been absent. The goal of this study was to provide some empirical data concerning the effect on students' achievement in mathematics through writing comment on students' homework papers. Specifically, the problem was to determine if teachers' comments on mathematics homework papers cause significant improvement in students' mathematics achievement.

The hypotheses for this study, in condensed form, stated the following: teacher comments on mathematics homework papers will significantly increase students' achievement in fourth, sixth and ninth grades. There will be no significant grade effect, and there will be no significant interaction between the treatment effect and grade levels.

Design of the Study

Six fourth grade mathematics teachers, six sixth grade mathematics teachers, and six ninth grade mathematics teachers were randomly selected from all fourth, sixth and ninth grade mathematics teachers in the city of Mecca, to participate in the study. For each selected teacher, two classes were selected from the eligible classes. One class was assigned to the experimental group and one was assigned to the control group. Pre- and post-tests were administered to both groups.

The Split-Plot design was selected for the study in order to measure the treatment effect, the grade level effect and the interaction between the treatment effect and the grade level. The sample consisted of 1,060 students enrolled in 36 classes. Twelve classes in each grade level. Six classes were assigned as the experimental group and the other six were assigned as the control group.

For ten weeks the teachers collected homework papers three times weekly. The homework papers in the No-Comment group (Control group) were graded by simply marking problems right or wrong and giving the papers a grade. Homework papers in the Comment group (Experimental group) were graded by marking problems right or wrong and giving the papers a grade. Papers in the Experimental group had teachers' comments. If a student missed a problem the teacher marked where the error occurred and either gave or indicated how to find a correct solution. Each paper in this group had one or more written comments, free comments, whatever free comments the teacher chose to make.

Data consisted of pre-test and post-test scores for each student. Equivalent forms of the course achievement examination were used to measure students' achievement. The tests were administered in each classroom at a time selected by the teachers.

Analysis of variance was used to test the major and the minor hypotheses. For the minor hypotheses, additional data were required. The investigator obtained all the students' scores in the previous term mathematics examination. The students were divided into three groups for the minor hypotheses:

Group 1: All the students who scored less than 55 points in the

previous term mathematics examination.

Group 2: All the students who scored between 55 and 75 points in the previous term mathematics examination.

Group 3: All the students who scored more than 75 points in the previous term mathematics examination.

The aim of this grouping was to determine whether the treatment was ability dependent.

Result of the Data Analysis

As previously indicated, there was no significant grade level effect at the 0.05 level. There was no significant interaction between the treatment and grade level. Similar results were obtained for the minor hypotheses at the 0.05 level. The F-value for the treatment effect was significant at the 0.05 level.

The test of achievement hypothesis showed that for these data teachers' comments significantly increased students' achievement. The same results were obtained for the minor treatment hypotheses. The test of achievement for the minor hypotheses showed that teachers' comments

significantly (at the 0.05 level) increased student's achievement in each group.

CONCLUSIONS

The experiment described in the preceding chapters provided evidence relating to the question of whether teachers' comments on students' mathematics homework papers would improve student achievement. The following conclusions were drawn from this study.

1. The students in the experimental group (free comment) achieved higher scores than the students in the control group (no-comment). The overall treatment effect was significant at the 0.05 level.

2. Although unsupported professional opinion seemed to hold that older students would be less responsive than younger ones, such was not the experimental conclusion. Rather, the evidence seemed to support a belief in the same responsiveness among students in fourth, sixth and ninth grades.

Grade level did not influence students' performance. The effect of grade level was not significant at the 0.05 level.

3. While free comment method is a more effective method than no-comment, it equally improved students' performance in fourth, sixth and ninth grade mathematics

classes. Type of correcting students' homework progress did not interact with students' grade levels. There is no significant interaction between the treatment effect and students' grade level.

4. As previously indicated, the aim of dividing the students into three different groups was to determine whether the treatment was ability dependent. The evidence seemed to support a belief of the same effectiveness among students of low, medium and high ability.

Three studies investigated teachers' comments incorporated into the classroom structure in a slightly different way. Page (1958) used 2,189 students in grades 7 through 12. Page randomly assigned students to one of the three groups in every class. Teachers wrote no comment on tests for students in one group, and they wrote standardized comments depending on the letter grade on the test for the students in another group. Teachers wrote free comments on student tests in the third group. Free comment students achieved higher scores than specified comment students, and specified comment did better than no-comments. All differences, except between free and specified comments were significant. The overall treatment effect, furthermore, was significant at the 0.01 level.

Austin (1976) made a similar study using 222

students in nine classes grades 4, 9 and 10. In each class, students were matched and randomly assigned to a Comment or No-Comment group. After six weeks the two groups were compared using teacher-prepared examinations. A significant difference favoring the comment group was found in two of nine classes, when all students were pooled. Comments improved students' achievement in geometry and one of two general mathematics classes and had no effect on algebra in one fourth-grade class.

Schoen and Krey (1974) also made a study involving written comments on homework for twelve weeks, several methods of writing comment were compared using 147 pre-service elementary education majors. No differences were found on test scores as a result of teachers' comments.

The last study (Schoen and Krey) used college students and a large class size.

These studies, as well as the others reviewed, differed from this study in several respects. This study involved teacher comments on homework papers rather than on test papers. The students did homework at home rather than in the classrooms as tests. The homework assignments were done in a non-competitive atmosphere. The students enrolled in this study were males. They attended schools for boys only, and all the teachers were males.

Results from this study provided needed empirical data regarding teacher comments on homework papers in mathematics, at elementary and secondary level. These data revealed that teacher comments significantly increase achievement in mathematics for students in the fourth, sixth and ninth grades in Saudi Arabia. Furthermore, the grade level nor ability level did not significantly affect the difference in student achievement.

RECOMMENDATIONS

The recommendations which are set forth in this section have been drawn from the results and conclusions of the study. These recommendations are of two types: recommendations for further research, and recommendations for educational practice in Saudi Arabia.

RECOMMENDATIONS FOR FURTHER RESEARCH

1. Replication of this study in both mathematics and other disciplines should be conducted to test the generalizability of the results and conclusions of this study.

2. Replication of this study in both the same grade levels and different grade levels should be conducted.

3. The experimental treatment used in this study

lasted ten weeks. Research for longer time periods should be conducted.

4. Replication of this study in boys' schools and girls' schools should be conducted to test the generalizability of the results and conclusions of this study.

5. Research should be conducted to determine how teachers can most effectively use the comment method in the classroom.

6. As was indicated in this study, the words of free comments were not chosen at random. It could not, therefore, be demonstrated what type of free comments had more effect. Another study might identify a population of students for whom a particular type of remark might be appropriate, and then randomly select subjects and controls from the population. In such a way the words themselves would be brought under truly experimental observation.

RECOMMENDATIONS FOR EDUCATIONAL PRACTICE IN SAUDI ARABIA

Any generalization based on the findings of this study must take into account the limitations of the study. With these limitations in mind, it is the investigator's opinion that, pending further research such as that described above, the following are valid recommendations for educational practice.

1. Fourth grade, sixth grade and ninth grade mathematics teachers should write free comment on students' homework papers.

2. Teachers' workshops should be conducted in order to:

a) Acquaint teachers with the rational and use of the free comment technique.

b) Promote the use of free comment method.

3. Teachers should be assigned appropriate teaching load in order to have enough time to write comments on students' mathematics homework papers.

Teachers' comments are one form of teacher-produced motivation accessible to nearly every teacher. When the average teacher takes the time and trouble to write such comments (believed to be encouraging) on student papers, these apparently have a measurable and potent effect upon student effort, or attention, or attitude, or whatever it is which causes learning to improve. A model has been designed and tested for measuring this motivation and variables related to it. Countless possibilities for future exploration remain a challenge to the serious student of student motivation.

BIBLIOGRAPHY

Anderson, W. E. An attempt through the use of experimental techniques to determine the effect of home assignments upon scholastic success. Journal of Educational Research. 1946, 40, 141-143.

Austin, J. D. and Austin, K. A. Homework grading procedures in junior high mathematics classes. School Science and Mathematics. 1974, 74, 269-272.

Austin, J. D. Do comments on mathematics homework affect students' achievement? School Science and Mathematics, 1976, 76, 159-164.

Bradely, R. M. An experimental study of individualized versus blanket type homework assignments in elementary school mathematics. Dissertation Abstract, XXVII, 3874-A.

Brooks, E. C. The value of home study under parental supervision. Elementary School Journal, 1916, 17, 187-194.

Crawford, C. C. and Carmichael, J. A. The value of home study. Elementary School Journal, 1937, 38, 194-200.

Di Napoli, P. Homework in the New York City elementary schools. Teacher College, Columbia University Contribution to Education No. 719, 1937, Bureau of Publications, Columbia University.

Foran, T. G. and Weber, Sister M. M. An experimental study of the relation of homework to achievement in arithmetic. Mathematics Teacher, 1939, 32, 212-214.

Goldstein, A. Does homework help? A review of research. Elementary School Journal, 1960, I, 212-224.

Gray, R. F. and Allison, D. E. An experimental study of relationship of homework to pupil success in computation with fractions. School Science and Mathematics, 1971, 71, 339-346.

Hines, V. A. Homework and achievement in plane geometry. The Mathematics Teacher, 1957, 50, 27-29.

Hudson, J. A. A pilot study of the influence of homework in seventh grade mathematics and attitudes toward homework in the Fayetteville public schools. Dissertation Abstracts, XXVI, 906.

Johnston, J. O. When to correct arithmetic problems--a critical variable. School Science and Mathematics, 1969, 69, 799-805.

Kock, E. A. Jr. Homework in arithmetic. The Arithmetic Teacher, 1965, 12, 9-13.

Leonard, Sister Mary Hilda. An experimental study of homework at the intermediate grade level. Dissertation Abstract, XXVI, 3782.

Maertens, N. An analysis of the effects of arithmetic homework upon the arithmetic achievement of third grade pupils. The Arithmetic Teacher, 1969, 16, 383-389.

Maertens, N. and Johnston, J. Effects of arithmetic homework upon the attitudes and achievement of fourth, fifth and sixth grade pupils. School Science and Mathematics, 1972, 72, 117-126.

Mason, L. C. A study of the influence of compulsory homework on the achievement of college students in college algebra. Dissertation Abstracts, XXVIII, 65-A.

Montgomery, C. F. An evaluation of required home study in junior high school arithmetic and English. Unpublished Master's thesis, University of West Virginia, 1933.

Mulry, J. G. We need research on homework. National Education Association Journal, L (April) 1961, 49.

Nietling, L. C. Using problems to initiate the study of certain topics in mathematics. Unpublished Ph.D. Dissertation, The Ohio State University, Columbus, Ohio, 1968.

Page, E. B. Teacher comments and student performance: a seventy-four classroom experiment in school motivation. The Journal of Education Psychology, 1958, 49, 178-181.

Schoen, Harold L. and Kreye, Bettibel, C. Five forms of written feedback to homework in a mathematics course for elementary teachers. Journal of Research in Mathematics Education, 1974, 5, 140-145.

Skinner, B. F. Science and human behavior. New York, Macmillan, 1953.

Small, D. E. A study of two methods of checking homework in high school geometry class. The Mathematics Teacher, 1967, 60, 149-152.

Steiner, M. A. Value of home-study assignments. School and Society, 1934 (July), 30, 20-24.

Swenson, Esther J., Cost, James C., and Taylor, G. Y. Research on homework. Journal of Education, CXXXVII (March) 1955, 20-22.

Teahan, E. G. Required home study is unwise. American School Board Journal, 1935 (November), 41, 41.

Ten Brinke, D. P. Homework: an experimental evaluation of the effect on achievement in mathematics in grades seven and eight. Dissertation Abstracts, XXVII, 4176-A.

Whelan, Joseph Arthur. An analysis of the effect of systematic homework in two fourth grade subjects. Dissertation Abstract XXVI, 5143.

Zastsow, M. C. The role of homework in a college mathematics course as related to the frequency principle in learning process. Dissertation Abstracts, XXVIII, 2153-A-2154-A.

APPENDIX A

APPENDIX A

RECRUITMENT LETTER TO TEACHERS

Dear Teacher:

What is the purpose of the study?

As every teacher knows, much of the talk about motivation is pretty foggy. No one seems to know enough to help newer teachers; what may be needed is more verbalizable principles which will save newcomers to teaching, and their students. Too often the people talking about motivation (i.e., what makes students perform in a desirable way) are either talking through their hats, saying what no one knows because no one has tested it, or basing their talk on laboratory experiments far removed from public schools and our problems there. Consequently, we often have different ideas without the opportunity to test them in any scientific way.

This lack of authoritative information is well exemplified by the subject of written comments on papers being returned to students. Should one comment on papers beyond the letter grade? Some teachers write faithfully on every paper. Other teachers never write comments, regarding them as a waste of time and effort. Still others do not write comments but feel vaguely guilty about it, as if they were somehow slacking their jobs.

And still other teachers write comments only on the very good papers, or very poor ones.

Similarly, teachers differ in the types of comment. One teacher will have blanket comments, e. g., "Good," which he will administer to papers that achieve a certain level. Another will simply point out strengths or, more often, weaknesses within the paper itself.

Whatever your own practice, you obviously do it for one of two reasons: you believe that your practice is the right practice and will produce some desirable effect on student's performance; or you believe you should do something else, something more thorough, but the limitation of your teaching time prevents it.

Who is right? If there were a "right" answer for everyone, or for the average teacher, then we might be able to advise the new teacher in a meaningful way. But perhaps there is a different answer for different teachers, or for different subjects, or different grade levels, or different students. If so, what are those differences? And how can you tell which students will react in which way? Are there no answers? No laws? None have been given. Therefore this study aims to explore for such answers in a much bigger way, and one much closer to your actual classroom procedures than has ever been done before.

Who are the subjects of the experiment? The

subjects of the study are around 1,000 students of public elementary and secondary schools within the city of Mecca. We are investigating the question: Do teacher comments on mathematics homework affect student achievement? And by "achievement" we mean what most teachers mean: learn more in our classrooms than they might otherwise. This is, after all, the way we commonly judge student achievement in our courses: how they perform on tests which we ourselves make and administer.

We are asking, therefore, that you proceed in your normal way, teaching what you are teaching in the way that you prefer. But we ask one change only: that when you are returning homework papers to the students you write comments on homework papers to the students in the experimental group only and write no comments at all on homework papers to the students in the control group.

Will there be a big difference between the achievement of these groups on the second test? In your one class, probably not, except that caused by chance. But when a great number of us pool our results, and treat the results in a certain way, the effects might turn out to be large, visible, and reliable. And we may all learn a tremendous amount from the pooling which would be impossible in any other way.

Suggestions

1. Please forget you are going to report the words to anyone else. Be natural. Take up the stack of free comment papers and go through them in any fashion you want, only finding something to say on each. Remember, when the study is published, no comment will be attached to any particular teacher or particular student and no names will be used.

2. Write anything that occurs to you in the circumstances. There is not any "right" or "wrong" comment for this study. A comment is "right" for the study if it conforms with your own feelings and practices. Here, for example, are some freedoms you have:

Address the student by name or not, just as it occurs to you.

Make your comment general or direct to one part of the homework, just as occurs.

Make your comment friendly or formal.

Make it about this homework or about the student's work in general.

3. Teachers should address themselves to a child's situation but not to judge his character and personality.

4. Teachers are advised to avoid diagnosis and

prognosis in dealing with students. Diagnosing children is dangerous. Labeling is disabling. The diagnosis may become the disease. A child often lives up to a teacher's negative prediction. He becomes what he is told he is. It is damaging to tell a child where he will end up. Destination may become destinies. Doom predictions create psychological fissures in the lives of children. Generations of children, especially minority children, left school because they have been convinced of their intellectual anemia by destructive persuasion.

5. Helpful correction is direction. It describes processes. It does not judge products or persons. Children need guidance, not criticism. Communication, like health, depends on acts of prevention. An enlightened teacher learns to omit messages that make a child feel foolish, guilty, and vengeful. He deliberately avoids questions and comments that are likely to incite resentment and invite resistence.

It is said that gentlemen never insult unintentionally. Some teachers do. They hurt children inadvertently. A teacher with an acid tongue is a health hazard. His caustic comments deflate self esteem and block learning. Hurt children grow preoccupied with revenge fantasies.

The following remarks were made by teachers almost

without awareness of their tragic impact:

> "You are relying on your own judgment again. Believe me, it is a poor guide."
>
> "Your intelligence is not good enough for this class. Why don't you transfer to a school more commensurate with your disabilities?"
>
> "You don't need a psychologist; you need a vacuum cleaner. Your mind is cluttered with junk."

There is no place for devastating remarks in teacher-child communication. A professional teacher shuns comments that casually destroy a child's self-esteem. A teacher's role is to heal, not to injure. A teacher with a critical disposition and a gifted tongue has a grave responsibility. He must protect young children from his deadly talent, either by learning new ways of communicating or by choosing another calling. Teachers are often unaware of the impact of their words on the lives of children. Do enlightened methods of communication make a difference? Can children tell the difference between benign and destructive messages? Do they respond differently to them? The following incident by sixth-grader answers these questions:

> The same kids who are cooperative in Mr. A's class, go off like firecrackers in Mr. B's class. The way a child behaves depends on the teacher. For instance, when Mohammed returned to school after a week of absence, Mr. A said "Welcome back. We missed you." The child was glad and behaved well. Mr. B said "No wonder it was so quiet last week. Mohammed was out." Mohammed did not stop making noise for the rest of the hour.

Congruent communication* can transform education. It strikes not just at the trapping of teaching, but at the hearts of learning. Yet it has not been tried in our schools. Never has its full force been lavished on children to enrich their personalities and enable their lives.

Congruent communication is an achievement. It requires learning and rehearsing and self-discipline. It is not just "doing what comes naturally." Like all skills, it demands practice. Like all acts, it requires selection. It is consoling, but untrue to claim that in good relationships one can say anything with impunity. It is like believing that when in good health one can swallow anything without harm--including poison.

---

*Communication that is harmonious, authentic, where words fit feeling.

APPENDIX B

Analysis of Variance
For All the Students
By Grade Level

Variable for Entire Population	Sum 61395	Mean 57.92	Std.Dev. 11.63	Variance 135.19	N 1060
Fourth Grade	21596	57.90	11.96	143.07	373
Experimental Group	11749	63.85	10.42	108.57	184
Control Group	9847	52.10	10.44	108.92	189
Sixth Grade	19109	58.44	11.36	129.10	327
Experimental Group	10460	67.39	11.19	125.30	165
Control Group	8649	53.39	9.11	82.95	162
Ninth Grade	20690	57.47	11.53	132.87	360
Experimental Group	11095	62.33	11.28	127.16	178
Control Group	9595	52.72	9.66	93.25	182

Analysis of Variance
For All the Students

Variable for Entire Population	Sum 61395	Mean 57.92	Std.Dev. 11.63	Variance 135.19	N 1060
Experimental Group	33304	63.19	10.96	120.05	527
Control Group	28091	52.70	9.78	95.60	533

The following tables display the distribution of all the students in each class. The total number of the students was 1,060.

Fourth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	22	35	37	35	25	30	184
Control Group	23	35	40	34	25	32	189
Total	45	70	77	69	50	62	373

Sixth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	26	33	21	25	33	27	165
Control Group	22	35	23	25	32	25	162
Total	48	68	44	50	65	52	327

Ninth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	33	26	35	32	27	25	178
Control Group	32	27	36	33	25	29	182
Total	65	53	71	65	52	54	360

Analysis of Variance
For the First Group
By Grade Level

Variable for Entire Population	Sum 27548	Mean 50.92	Std.Dev. 9.74	Variance 94.86	N 541
Fourth Grade	8872	50.41	10.65	113.47	176
Experimental Group	4645	56.65	8.96	80.28	82
Control Group	4227	44.97	8.99	79.36	94
Sixth Grade	8926	52.20	9.49	90.11	171
Experimental Group	5140	57.11	9.60	92.12	90
Control Group	3786	46.74	5.63	31.69	81
Ninth Grade	9750	50.26	9.00	81.02	194
Experimental Group	5140	54.68	8.88	78.93	94
Control Group	4610	46.10	6.91	47.77	100

Summary of Analysis of Variance
For Group 1

Variable for Entire Population	Sum 27548	Mean 50.92	Std.Dev. 9.74	Variance 94.86	N 541
Experimental Group	14925	56.11	9.18	84.33	266
Control Group	12623	45.90	7.35	53.97	275

Group 1

The total number of the students who scored up to 55 points in the previous term examinations was 541. The tables below display the sample in detail.

Fourth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	9	17	21	11	12	12	82
Control Group	7	18	24	20	14	11	94
Total	16	35	45	31	26	23	176

Sixth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	17	17	13	11	19	13	90
Control Group	13	16	15	13	12	12	81
Total	30	33	28	24	31	25	171

Ninth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	18	12	26	16	12	10	94
Control Group	21	16	25	15	13	10	100
Total	39	28	51	31	25	20	194

Analysis of Variance
For Group 2
By Grade Level

Variable for Entire Population	Sum	Mean	Std.Dev.	Variance	N
	23612	63.13	7.36	54.25	374
Fourth Grade	7289	61.77	7.53	56.67	118
Experimental Group	3904	66.17	6.60	43.62	59
Control Group	3385	57.37	5.50	31.34	59
Sixth Grade	7388	63.14	7.86	61.87	117
Experimental Group	3705	68.61	6.40	40.96	54
Control Group	3683	58.46	5.69	32.41	63
Ninth Grade	8935	64.28	6.60	43.68	139
Experimental Group	4895	68.94	4.62	21.37	71
Control Group	4040	59.41	4.53	20.54	68

Summary of Analysis of Variance
For Group 2

Variable for Entire Population	Sum	Mean	Std.Dev.	Variance	N
	23612	63.13	7.36	54.25	374
Experimental Group	12504	67.96	5.95	35.40	184
Control Group	11108	58.46	5.31	28.23	190

Group 2

The following tables display in detail the distribution of the students who scored between 55 and 75 points in the previous term examination. The total number for this group was 374.

Fourth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	4	13	10	15	5	12	59
Control Group	8	12	11	10	6	12	59
Total	12	25	21	25	11	24	118

Sixth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	6	10	4	13	9	12	54
Control Group	7	16	8	8	13	11	63
Total	13	26	12	21	22	23	117

Ninth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	14	14	6	10	14	13	71
Control Group	11	11	6	16	9	15	68
Total	25	25	12	26	23	28	139

Analysis of Variance
For Group 3
By Grade Level

Variable for Entire Population	Sum 10235	Mean 70.59	Std.Dev. 9.04	Variance 81.77	N 145
Fourth Grade	5435	68.80	8.56	73.21	79
Experimental Group	3200	74.42	5.90	34.77	43
Control Group	2235	62.08	6.02	36.25	36
Sixth Grade	2795	71.67	9.62	92.54	39
Experimental Group	1615	76.90	8.13	66.19	21
Control Group	1180	65.55	7.45	55.55	18
Ninth Grade	2005	74.26	8.51	72.51	27
Experimental Group	1060	81.54	4.74	22.43	13
Control Group	945	67.50	4.70	22.11	14

Summary of Analysis of Variance
For Group 3

Variable for Entire Population	Sum 10235	Mean 70.59	Std.Dev. 9.04	Variance 81.77	N 145
Experimental Group	5875	76.30	6.85	46.97	77
Control Group	4360	64.12	6.52	42.49	68

Group 3

The total number of students who scored more than 75 points in the previous term examination was 145. The following tables display in detail the distribution of these students.

Fourth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	9	5	6	9	8	6	43
Control Group	8	5	5	4	5	9	36
Total	17	10	11	13	13	15	79

Sixth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	3	6	4	1	5	2	21
Control Group	2	3	-	4	7	2	18
Total	5	9	4	5	12	4	39

Ninth Grade

Teacher	1	2	3	4	5	6	Total
Experimental Group	1	-	3	6	1	2	13
Control Group	-	-	5	2	3	4	14
Total	1	-	8	8	4	6	27

Analysis of Variance
For Entire Population (Fourth Grade)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Fourth Grade	21596	57.90	11.96	143.07	373
Experimental Group	11749	63.85	10.42	108.57	184
Teacher 1	1405	63.86	6.35	40.31	22
2	2164	61.83	9.41	88.50	35
3	2280	61.62	9.79	95.91	37
4	2290	65.43	10.10	102.02	35
5	1610	64.40	12.13	146.50	25
6	2000	66.67	13.02	169.54	30
Control Group	9847	52.10	10.44	108.92	189
Teacher 1	1235	53.69	9.91	98.22	23
2	1715	49.00	10.56	111.47	35
3	2032	50.80	10.01	100.27	40
4	1800	52.94	8.80	77.45	34
5	1285	51.40	13.03	169.83	25
6	1780	55.62	9.98	99.60	32

Analysis of Variance
For Entire Population (Sixth Grade)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Sixth Grade	19109	58.44	11.36	129.10	327
Experimental Group	10460	63.40	11.19	125.30	165
Teacher 1	1610	61.92	11.84	140.15	26
2	2070	62.73	11.53	132.95	33
3	1275	60.71	10.99	120.71	21
4	1575	63.00	10.61	112.50	25
5	2190	66.36	10.40	108.24	33
6	1740	64.44	11.87	141.02	27
Control Group	8649	53.39	9.11	82.95	162
Teacher 1	1135	51.59	7.77	60.44	22
2	1814	51.83	8.95	80.09	35
3	1180	51.30	8.29	68.67	23
4	1335	53.40	9.54	91.08	25
5	1820	56.87	10.06	101.21	32
6	1365	54.60	8.77	76.92	25

Analysis of Variance
For Entire Population (Ninth Grade)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Ninth Grade	20690	57.47	11.53	132.87	360
Experimental Group	11095	62.33	11.28	127.16	178
Teacher 1	2005	60.76	9.93	98.63	33
2	1585	60.96	9.27	86.04	26
3	2095	59.86	11.28	127.19	35
4	2060	64.37	15.44	238.31	32
5	1710	63.33	9.20	84.61	27
6	1640	65.60	10.24	104.83	25
Control Group	9595	52.72	9.66	93.25	182
Teacher 1	1635	51.09	8.77	76.99	32
2	1375	50.92	9.10	82.76	27
3	1825	50.69	12.54	157.36	36
4	1810	54.85	7.55	57.01	33
5	1335	53.40	9.97	99.42	25
6	1615	55.69	8.21	67.36	29

Analysis of Variance
For Combined Population
By Teachers

Variable for Entire Population	Sum	Mean	Std.Dev.	Variance	N
	61395	57.92	11.63	135.19	1060
Fourth Grade	21596	57.90	11.96	143.07	373
Teacher 1	2640	58.67	9.73	94.77	45
2	3879	55.41	11.84	140.27	70
3	4312	56.00	11.25	126.50	77
4	4090	59.27	11.32	128.14	69
5	2895	57.90	14.07	198.05	50
6	3780	60.97	12.73	162.16	62
Sixth Grade	19109	58.44	11.36	129.10	327
Teacher 1	2745	57.19	11.34	128.62	48
2	3884	57.12	11.59	134.25	68
3	2455	55.80	10.67	113.89	44
4	2910	58.20	11.10	123.22	50
5	4010	61.69	11.22	125.99	65
6	3105	59.71	11.52	132.76	52
Ninth Grade	20690	57.47	11.53	132.87	360
Teacher 1	3640	56.00	10.50	110.31	65
2	2960	55.85	10.41	108.40	53
3	3920	55.21	12.72	161.74	71
4	3870	59.54	12.92	166.97	65
5	3045	58.56	10.72	115.03	52
6	3255	60.28	10.39	107.94	54

Analysis of Variance
For Group 1 (Fourth Grade)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Fourth Grade	8872	50.41	10.65	113.47	176
Experimental Group	4645	56.65	8.96	80.28	82
Teacher 1	550	61.11	4.17	17.36	9
2	945	55.59	8.27	68.38	17
3	1195	56.90	8.58	73.69	21
4	600	54.54	10.11	102.27	11
5	660	55.00	9.04	81.82	12
6	695	57.42	11.96	142.99	12
Control Group	4227	44.97	8.91	79.36	94
Teacher 1	300	42.86	8.10	65.48	7
2	745	41.39	9.36	87.66	18
3	1102	45.97	9.17	84.17	24
4	1010	50.50	8.41	70.79	20
5	595	42.50	8.03	64.42	14
6	475	43.18	6.03	36.36	11

Analysis of Variance
For Group 1 (Sixth Grade)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Sixth Grade	8926	52.20	9.49	90.11	171
Experimental Group	5140	57.11	9.60	92.12	90
Teacher 1	960	56.47	10.42	108.64	17
2	940	55.29	8.56	73.34	17
3	720	55.38	9.89	97.76	13
4	620	56.36	9.77	95.45	11
5	1160	61.05	7.37	54.39	19
6	740	56.92	12.17	148.08	13
Control Group	3786	46.74	5.63	31.69	81
Teacher 1	605	46.54	5.55	30.77	13
2	721	45.06	5.66	32.06	16
3	710	47.33	5.94	35.24	15
4	620	47.69	6.33	40.06	13
5	570	47.50	5.43	29.54	12
6	560	46.67	5.36	28.79	12

Analysis of Variance
For Group 1 (Ninth Grade)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Ninth Grade	9750	50.26	9.00	81.02	194
Experimental Group	5140	54.68	8.88	78.93	94
Teacher 1	970	53.89	7.77	60.46	18
2	640	53.33	6.85	46.97	12
3	1445	55.58	8.52	72.65	26
4	840	52.50	11.40	130.00	16
5	675	56.25	8.82	77.84	12
6	570	57.00	10.33	106.67	10
Control Group	4610	46.10	6.91	47.77	100
Teacher 1	980	46.67	6.58	43.33	21
2	735	45.94	7.79	60.73	16
3	1095	43.80	6.96	48.50	25
4	730	48.67	4.80	23.09	15
5	605	46.54	8.75	76.60	13
6	465	46.50	5.80	33.61	10

Analysis of Variance
For Combined Population (Group 1)
By Teacher

Variable for Entire Population	Sum 27548	Mean 50.92	Std.Dev. 9.74	Variance 94.86	N 541
Fourth Grade	8872	50.41	10.65	113.47	176
Teacher 1	850	53.12	11.09	122.92	16
2	1690	48.28	11.31	127.86	35
3	2297	51.04	10.40	108.22	45
4	1610	51.93	9.10	82.80	31
5	1255	48.27	10.48	109.88	26
6	1170	50.87	12.03	144.66	23
Sixth Grade	8926	52.20	9.49	90.11	171
Teacher 1	1565	52.17	9.88	94.73	30
2	1661	50.33	8.87	78.67	33
3	1430	51.07	8.86	78.44	28
4	1240	51.67	9.05	81.88	24
5	1730	55.81	9.41	88.49	31
6	1300	52.00	10.70	114.58	25
Ninth Grade	9750	50.26	9.00	81.02	194
Teacher 1	1950	50.00	7.95	63.16	39
2	1375	49.11	8.17	66.76	28
3	2540	49.80	9.74	94.96	51
4	1570	50.64	8.92	79.57	31
5	1280	51.20	9.92	98.50	25
6	1035	51.75	9.77	95.46	20

Analysis of Variance
For Group 2 (Fourth Grade)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Ninth Grade	7289	61.77	7.53	56.67	118
Experimental Group	3904	66.17	6.60	43.62	59
Teacher 1	230	57.50	2.89	8.33	4
2	849	65.31	4.87	23.73	13
3	630	63.00	4.83	23.33	10
4	1025	68.33	4.50	20.24	15
5	350	70.00	7.07	50.00	5
6	820	68.33	9.13	83.33	12
Control Group	3385	57.37	5.60	31.34	59
Teacher 1	440	55.00	7.07	50.00	8
2	675	56.25	3.11	9.66	12
3	605	55.00	3.87	15.00	11
4	575	57.50	7.55	56.94	10
5	355	59.17	8.01	64.17	6
6	735	61.25	2.26	5.11	12

Analysis of Variance
For Group 2 (Sixth Grade)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Sixth Grade	7388	63.4	7.86	61.87	117
Experimental Group	3705	68.61	6.40	40.96	54
Teacher 1	415	69.17	3.76	14.17	6
2	650	65.00	4.71	22.22	10
3	270	67.50	6.45	41.67	4
4	885	68.08	8.55	73.08	13
5	640	71.11	6.01	36.11	9
6	845	70.42	5.82	33.90	12
Control Group	3683	58.46	5.69	32.41	63
Teacher 1	405	57.86	2.67	7.14	7
2	908	56.75	7.14	51.00	16
3	470	58.75	6.94	48.21	8
4	445	55.62	6.78	45.98	8
5	780	60.00	4.56	20.83	13
6	675	61.36	2.33	5.45	11

Analysis of Variance
For Group 2 (Ninth Grade)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Ninth Grade	8935	64.28	6.61	43.68	139
Experimental Group	4895	68.94	4.62	21.37	71
Teacher 1	960	68.57	4.13	17.03	14
2	945	67.50	5.09	25.96	14
3	405	67.50	6.89	47.50	6
4	710	71.00	4.59	21.11	10
5	955	68.21	3.17	10.03	14
6	920	70.77	4.49	20.19	13
Control Group	4040	59.41	4.53	20.54	68
Teacher 1	655	59.54	5.68	32.27	11
2	640	58.18	5.13	26.36	11
3	370	61.67	4.08	16.67	6
4	945	59.06	4.55	20.73	16
5	540	60.00	4.33	18.75	9
6	890	59.33	3.72	13.81	15

Analysis of Variance
For Combined Population (Group 2)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Entire Population	3612	63.13	7.36	54.25	374
Fourth Grade	7289	61.77	7.53	56.67	118
Teacher 1	670	55.83	5.97	35.61	12
2	1524	60.96	6.13	37.62	25
3	1235	58.81	5.89	34.76	21
4	1600	64.00	7.90	62.50	25
5	705	64.09	9.17	84.09	11
6	1555	64.79	7.44	55.39	24
Sixth Grade	7388	63.14	7.86	61.87	117
Teacher 1	820	63.08	6.63	43.91	13
2	1585	59.92	7.44	55.35	26
3	740	61.67	7.78	60.61	12
4	1330	63.33	9.92	98.33	21
5	1420	64.54	7.54	56.93	22
6	1520	66.09	6.39	40.81	23
Ninth Grade	8935	64.28	6.61	43.68	139
Teacher 1	1615	64.60	6.60	43.58	25
2	1585	63.40	6.88	47.33	25
3	775	64.58	6.20	38.45	12
4	1655	63.65	7.42	55.11	26
5	1495	65.00	5.43	29.54	23
6	1810	64.64	7.06	49.87	28

Analysis of Variance
For the Fourth Grade (Group 3)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Fourth Grade	5435	68.80	8.56	73.21	79
Experimental Group	3200	74.42	5.90	34.77	43
Teacher 1	625	69.44	4.64	21.53	9
2	370	74.00	4.18	17.50	5
3	455	75.83	3.76	14.17	6
4	665	73.89	3.33	11.11	9
5	600	75.00	6.55	42.86	8
6	485	80.83	7.36	54.17	6
Control Group	2235	62.08	6.02	36.25	36
Teacher 1	495	61.87	2.59	6.70	8
2	295	59.00	2.24	5.00	5
3	325	65.00	3.53	12.50	5
4	215	53.75	11.09	122.92	4
5	335	67.00	7.58	57.50	5
6	570	63.33	2.50	6.25	9

Analysis of Variance
For the Sixth Grade (Group 3)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Sixth Grade	2795	71.67	9.62	92.54	39
Experimental Group	1615	76.90	8.13	66.19	21
Teacher 1	235	78.33	5.77	33.33	3
2	480	80.00	4.47	20.00	6
3	285	71.25	6.29	39.58	4
4	70	70.00	0	0	1
5	390	78.00	13.96	195.00	5
6	155	77.50	3.53	12.50	2
Control Group	1180	65.55	7.45	55.55	18
Teacher 1	125	62.50	3.53	12.50	2
2	185	61.67	5.77	33.33	3
4	270	67.50	6.45	41.67	4
5	470	67.14	10.35	107.14	7
6	130	65.00	0	0	2

Analysis of Variance
For the Ninth Grade (Group 3)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Ninth Grade	2005	74.26	8.51	72.51	27
Experimental Group	1060	81.54	4.74	22.43	13
Teacher 1	75	75.00	0	0	1
3	245	81.67	2.89	8.33	3
4	510	85.00	0	0	6
5	80	80.00	0	0	1
6	150	75.00	7.07	50.00	2
Control Group	945	67.50	4.70	22.11	14
Teacher 3	360	72.00	4.47	20.00	5
4	135	67.50	3.53	12.50	2
5	190	63.33	2.89	8.33	3
6	260	65.00	0	0	4

Analysis of Variance
For Combined Population (Group 3)
By Teacher

Variable	Sum	Mean	Std.Dev.	Variance	N
Entire Population	10235	70.59	9.04	81.77	145
Fourth Grade	5435	68.80	8.56	73.21	79
Teacher 1	1120	65.88	5.37	28.86	17
2	665	66.50	8.51	72.50	10
3	780	70.91	6.64	44.09	11
4	880	67.69	11.48	131.73	13
5	935	71.92	7.78	60.58	13
6	1055	70.33	10.08	101.67	15
Sixth Grade	2795	71.67	9.62	92.54	39
Teacher 1	360	72.00	9.75	95.00	5
2	665	73.89	10.24	104.86	9
3	285	71.25	6.29	39.58	4
4	340	68.00	5.70	32.50	5
5	860	71.67	12.67	160.60	12
6	285	71.25	7.50	56.25	4
Ninth Grade	2005	74.26	8.51	72.51	27
Teacher 1	75	75.00	0	0	1
3	605	75.62	6.23	38.84	8
4	645	80.62	8.21	67.41	8
5	270	67.50	8.66	75.00	4
6	410	68.33	6.05	36.67	6

APPENDIX C

Interaction

Is an incident where the two independent variables (types of treatment and grades) act jointly to determine performance on the dependent variable (students' achievement).

Examples

1. If the comment method is more effective than the no-comment, and improves students' performance in one grade more than another grade, then we will have an interaction where there is no grade effect but there is a treatment effect.

2. If for one grade the comment method is more effective than the no-comment but for the other grades the no-comment is more effective than the comment method, then there will be an interaction with (no grade effect) and (no treatment effect).

There are no main effects in this example. Does this mean that neither of the two independent variables influenced the dependent variable? If there is no effect at all shouldn't the students improve equally under both methods in all grades? This indicates, then, that there should be an effect of the experimental treatment on the dependent variable. It is just that this effect is not covered by the influence of either of our main independent

variable alone, but rather, both of them are acting jointly to determine performance of the dependent variable. Notice that for one grade comment is more effective method than no-comment, but for the other grade, no-comment method is more effective than comment method. Here we have an example of an interaction in which neither main effect was significant.

Let us examine the last example in a slightly different way. Can we say anything about what will happen in our experiment if we know only the students' grade? Our answer to this question should be no, that is there is no grade effect. Can we say anything about what will happen in our experiment if we know only the correcting methods to which the students will be exposed? Your answer again should be no because there is no main effect of correcting method. What you need to know in order to say anything about the results of the experiment is whether the student is in the fourth, sixth or ninth grade and the type of correcting homework employed. The reason we need information on both of these variables is that, in this example, neither alone tells us anything about the student performance on the dependent variable; rather both of these variables act jointly to determine performance. This is what is meant by the concept of interaction. More specifically, it was said that for an

interaction to be observed one type of relationship held at one level of variable and another type of relationship held at another level of that variable. That is, for one grade, the comment method is more effective--this is one relationship--whereas for other grade, no-comment is more effective--this is yet another relationship. When these relationships are not parallel, we say that we have an interaction.

It is time now to point out in a more precise way what is meant by the term. Interaction is a term borrowed from the field of statistics. The statisticians' definition of an interaction rests upon whether or not the relationships or curves are indeed parallel. We all know what it means to have two parallel lines. Thus for an interaction to be significant, the lines must significantly depart from a parallel relation.

APPENDIX D

Sample of Teachers' Comments

I like the way you solved problem 3.

You got the right answer, but there is another way to solve the problem. Think about it.

I am glad to see your improvement. Nice work. Keep it that way.

I can see that you understand the multiplication table.

The method is right. You got a different answer because the error is in dividing 220 by 11.

Please notice the units you work with.

You need to work hard on the multiplication table.

In problem 2 you added yards to feet without converting to one unit.

APPENDIX E

Fourth Grade Test

Answer the following questions:

1) Find three different fractions equivalent to each of the following fractions:

   a) $\frac{1}{3}$ =

   b) $\frac{4}{7}$ =

   c) $\frac{3}{4}$ =

2) Arrange the following fractions starting with the smallest to the largest.

   a) $\frac{2}{5}$, $\frac{5}{6}$, $\frac{4}{5}$

   b) $\frac{2}{3}$, $\frac{1}{4}$, $\frac{2}{5}$

3) Select one of the following signs between the fractions, = , $<$, $>$

   a) $\frac{3}{5}$ ☐ $\frac{2}{5}$

   b) $\frac{1}{5}$ ☐ $\frac{1}{6}$

Sixth Grade Test

Answer the following questions:

1) The length of a string is

Yards	Feet	Inches
14	2	8

What is its length in inches?

2) A square piece of land 25 meters on a side needed to be fenced. The cost of a longitudinal meter was 250 Riyals. What would be the total cost of the fence?

3) A salesman sold all he had from a particular merchandise for 8400 Riyals. He made 15% profit on the sale price. What was the profit and for how much did he buy the merchandise?

Ninth Grade Test

Simplify the following:

1) $\frac{6X^2 + 25X + 25}{X^3 + 8} \times \frac{2X^2 + 9X + 10}{3X + 5}$

2) $\frac{X-2}{(X^2-1)} + \frac{10}{X^2-4X-5}$

3) $\frac{2X^3 + YX^2}{X-Y} \times \frac{4X^2-Y^2}{XY+X^2} + \frac{4X^3 + 4X^2Y + XY^2}{X + Y}$