DOCUMENT RESUME

ED 087 841 UD 014 079

TITLE	An Evaluation of Project for Increased Achievement and a New Outlook. State Urban Education Program.
INSTITUTION	New York Univ., N.Y. Center for Field Research and School Services.
SPONS AGENCY	New York City Board of Education, Brooklyn, N.Y.
PUB DATE	Jul 73
NOTE	72p.; Function Number 17-36474
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Compensatory Education Programs; Disadvantaged Youth; *Dropout Prevention; Economically Disadvantaged; Educationally Disadvantaged; Family School Relationship; Flexible Schedules; Guidance Programs; *High School Students; *Program Evaluation; Remedial Instruction; School Psychologists; Small Group Instruction; Urban Education
IDENTIFIERS	*New York City

ABSTRACT

The intent of this project was to improve the quality of the high school experience for a selected group of "high risk" freshmen by providing them with services beyond those normally made available. At each of six high schools, the guidance counselors together with the counselors of each school's feeder schools, selected a group of 250 incoming students whose records indicated severe learning disabilities that would make them higher potential dropouts. In general, these students who appeared economically as well as educationally disadvantaged, had to travel by bus to reach their schools which were situated in middle class economic settings. Among the additional services and modifications provided for these students were smaller classes, specially modified schedules, lighter course loads, special remedial instruction, intensive guidance and school-home liaison, and referrals to a school psychologist when needed. All six high schools were part of a simple project with an overall concern for supporting, guiding and enriching the school experience of an atypical group of students. Nevertheless, each school exercised a great deal of autonomy and operated a program unique to that school. In part these differences in program may have been directed by differences in the incoming student population. In part the differences in program may have been dicated by the balance or unbalance of strengths, values, and predilections among the project personnel involved. (Author/JM)

Function Number 17-36474

AN EVALUATION OF PROJECT FOR INCREASED ACHIEVEMENT AND A NEW OUTLOOK

STATE URBAN EDUCATION PROGRAM

An evaluation of a New York City school district educational project funded by the New York State Urban Education Program enacted at the 1970 Legislative Session of the New York State Legislature for the purpose of "meeting special educational needs associated with poverty." (Education Law 3602, subdivision 11 as amended.) Performed under contract with the Board of Education of the City of New York for the 1972–1973 school year.

Donald Payne
Robert Ruderman
Project Co-Directors

CENTER FOR EDUCATIONAL RESEARCH AND FIELD SERVICES
School of Education
New York University

July, 1973

New York University

School of Education
Center for Educational Research and Field Services
Office of Off-Campus Courses

51 Press Building
Washington Square
New York, N.Y. 10003
Telephone: (212) 598-2898, 3425

July 15, 1973

Dr. Anthony J. Polemeni
Director (Acting)
Bureau of Educational Research
Board of Education of the City of New York
110 Livingston Street - Room 724
Brooklyn, New York 11201

Dear Dr. Polemeni:

In fulfillment of the agreement dated June 11, 1973 between the New York City Public Schools and the Center for Educational Research and Field Services, I am pleased to submit three hundred copies of the final report, Project for Increased Achievement And A New Outlook.

The Bureau of Educational Research and the professional staff of the New York City Public Schools were most cooperative in providing data and facilitating the study in general. Although the objective of the team was to evaluate a project funded under State Urban Education, this report goes beyond this goal. Explicit in this report are recommendations for modifications and improvement of the program. Consequently, this report will serve its purpose best if it is studied and discussed by all who are concerned with education in New York City--the Board of Education, professional staff, students, parents, lay leaders, and other citizens. To this end, the study team is prepared to assist with the presentation and interpretation of its report. In addition, the study team looks forward to our continued affiliation with the New York City Public Schools.

You may be sure that New York University and its School of Education will maintain a continuing interest in the schools of New York City.

Respectfully submitted,

Arnold Spinner

Arnold Spinner
Director

AS:mh

TABLE OF CONTENTS

LIST OF TABLES

EXECUTIVE SUMMARY

Program Description

The intent of this project was to improve the quality of the high school experience for a selected group of "high risk" freshman by providing them with services beyond those normally made available. At each of six high schools, the guidance counselors together with the counselors of each school's feeder schools, selected a group of 250 incoming students whose records indicated "severe learning disabilities that make them higher potential drop-outs." In general, these students who appeared "economically as well as educationally disadvantaged," had to travel by bus to reach their schools which were situated in middle class economic settings. Among the additional services and modifications provided for these students were smaller classes, specially modified schedules, lighter course loads, special remedial instruction, intensive guidance and school-home liaison, and referrals to a school psychologist when needed. The staff designated to provide these services at each school included the following:

> 1 coordinator (full-time teacher); 1 grade advisor (full-time teacher); 1 full-time school psychologist (10 hours per week); and 1 family assistant (25 hours per week). The average estimated cost per student for these services was $240.

All six high schools were part of a simple project with an overall concern for supporting, guiding and enriching the school experience of an atypical group of students. Nevertheless each school exercised a great deal of autonomy and operated a program unique to that school. In part these differences in program may have been dictated by differences in the incoming student population. For example, at School A the average reading score of PIANO students was 4.94, while at School B the average was 7.35. In part the differences in program may have been dictated by the balance or unbalances of strengths, values, and predilections among the project personnel involved. For example, while the reading program at one school stressed specific skills, the program at another school gave more emphasis to thematic understanding and appreciation of literature. Procedurally, the students at one school attended PIANO only classes, at another PIANO remedial classes only, and at another were completely integrated into the regular school program. Thus at some schools there was an attempt at getting the students to achieve a strong identification with the PIANO project and at another school most of the PIANO students were unaware that they were part of a special program.

It should be noted in passing that these programmatic differences between schools are not reflected in the evaluation objectives nor for that matter, was any systematic attempt made to compare the possible effects of these differences.

Program Objectives

The proposal for PIANO stated the following program objectives:

1. To facilitate and ease the adjustment of disadvantaged alienated youth to high school.
2. To provide special supportive services to facilitate the student's awareness of his own potential and to help him develop that potential to the fullest extent.
3. To provide special guidance services for educational and vocational planning during and after high school.
4. To provide intensive remediation services to those students in the ninth year who need such services.
5. To open up and maintain a contact with the student's home and parent to facilitate resolution of problems which may be preventing the student from achieving.
6. To develop a climate in which the home and school can work together to create in the student a feeling of belonging to and identification with the school.
7. To improve attendance patterns to facilitate learning.

Evaluation Objectives

Roughly corresponding to these program objectives are the following objectives of evaluation which were prepared by the Bureau of Educational Research of the City of New York and appended to the PIANO proposal:

1. To determine whether at least 75% of the ninth grade students in each school in the program will improve by at least 8 school months in reading and arithmetic achievement from April 1972 to April 1973 on an appropriate standardized achievement test.
2. To determine whether at least 75% of the ninth grade students in the program in each school will improve by at least two scale points in appropriate school behavior from initial to final rating by the grade advisor.

3. To determine whether at least 75% of the ninth grade students in the program in each school will improve by at least two scale points in motivation and positive attitude toward school from initial to final rating by the grade advisor.
4. There will be a 30% increase in attendance of the ninth grade students in the program in each school.
5. There will be a decrease of at least 35% in disciplinary infraction of the ninth grade students in the program.
6. There will be an increase in home and parental contacts leading to a resolution of student referral problems.
7. 60% of the ninth grade students in the program in each school will pass all of their subjects in January 1973; 75% of the ninth grade students in the program in each school will pass all their subjects in June 1973.

In the next section, data related to each objective will be described and analyzed.

Recommendations

The recommendations in this section fall into two general categories, those concerned with evaluations of projects of this type, and those concerned with the design of future projects.

Concerning evaluations

1. It is strongly recommended that the personnel of new programs that are to be evaluated should be directly and seriously consulted in the development of evaluation objectives, criteria and measurement procedures. Not only might this result in better staff morale and attitude toward evaluation but it might also result in better evaluation designs. For example, had the project staff been consulted, they would have emphatically pointed out that in the original evaluation design it was completely unrealistic to base the evaluation of the changes in attitude and behavior of 250 students on the judgment of a single grade advisor. It is unlikely that a single person could have an adequate enough knowledge of such a large group of students to be able to rate them accurately and reliably.

2. Objectives should be supported by some rationale or hypotheses and should be based on past data when such data are available. For example, the expectation of the achievement of 8 months growth in as many months as measured by a standardized test normed on a national sample, may be unrealistic for a group of learners that has been gaining at a consistently slower rate over a period of many years. As previously discussed, objectives based on a historical regression approach may be more sensible in this situation.

Similarly, there probably exist extensive and relatively reliable attendance data for the students in experimental projects from which future trends, in terms of such factors as schools, seasons, and sex, could be predicted with relative accuracy. Changes in these trends due to some programmatic treatment could then be equated with "effect size" in experimental research. Hypotheses regarding the amount of change expected could then be offered in terms of the standard deviation of past distributions of attendance data. For example, a large change in attendance could be an increase of a single standard deviation over last year's average for some group. A small change might be a quarter of a standard deviation and so on. The point being that the objective would be stated in terms related to past data, current recommended research procedures, and based on some hypotheses or rationale for expecting an effect of a specific size.

3. Objectives should be sensitive to differences in goals and procedures at different project units. As discussed previously, different schools in the PIANO project operated fairly divergent programs and yet all were evaluated by the same criteria. Furthermore it is recommended that where there are differences among units, some systematic attempt be made to analyze the possible effects of these differences.

4. The persons who will actually implement the project evaluation should have the opportunity to interact with the project personnel at the very early stages of the project. This early involvement may be important for the following reasons:

a. there would be agreement on the precision of the data to be collected, times of test administrations and formats of records for recording and storing data, thus making the data collection procedures more efficient and less costly and allowing the data analysis procedures to be more powerful.

b. the improvement in efficiency and effectiveness of the data collection and analysis procedures could lead to interim or formative judgments which could be fed back into the project in time to guide its procedures on the basis of its current effectiveness or lack of it (e.g. if it was realized that the remedial reading program at School A, or the remedial math program at School C or the attendance rate at School D was significantly different than those at other schools, the "causes" of these differences might have been inferred and generalized to the programs at the other schools);

c. the formative or guiding effect of the evaluation data might have initiated a feeling among the staff that evaluation procedures were an integral and helpful part of the project rather than merely a summary judgment of how well they had done their jobs;

d. procedures for monitoring and describing the programs at each school, in particular their differences, could have been established to guide each unit in pursuing the uniqueness of its objectives.

Summary

In summary, we note that we cannot say, unequivocally, that the PIANO project accomplished *any* of the evaluations objectives set for it. In fact, this is the second consecutive year that evaluators have reached this judgment. Perhaps it is time to consider the alternative reasons for this state of affairs.

The first alternative that comes to mind is that the evaluation objectives are unrealistic. If this was simply an after the fact determination, this alternative might not merit serious consideration. However, there are certain aspects of the situation which *do* indicate that this alternative be seriously considered. For example, the PIANO evaluators for the previous year clearly stated their judgment that the objectives last year, which were similar to this year's, were "unrealistic." Either they "demanded too much growth of PIANO students" or they "did not tap the salient features of the programs." However, if anything, the objectives this year may have been even more demanding. In setting forth an expectation (as in objective one) that *at least 75%* of the students in PIANO will achieve at least 8 months growth during their 8 months in the Program before post-testing, an expectation is being set

forth that PIANO students will grow at a rate faster than the normal population. Based on the generally dismal record of the PIANO students' previous 8 school years this expectation does indeed seem unrealistic. In this regard predicted achievement based on individual historical regression seems a clear improvement, as previously discussed.

Beyond this on at least two different occasions early in the project, the PIANO staff sent written communications to the central office expressing their specific dissatisfactions with the original evaluation design. However, neither the staff, nor their objections were consulted in developing the evaluation design. For these reasons, past data, statements of previous evaluators, and the early expressions of staff dissatisfaction, a reasonable argument can be presented that the project was not being evaluated by fair or appropriate criteria.

Another alternative reason for the project's non-achievement may have to do with the relatively short period of time during which the project is expected to effect a significant change of behavior. For students with 8 school years of failure and 14 to 16 years of problematic existence, 8 school months may be a relatively short period of time to effect a change. Perhaps a more reasonable evaluation objective would consider a longer time in the project, i.e. two academic years, and a third and fourth year follow-up study.

A final alternative reason for lack of project achievement must of course be that the project is reaching too few students with any degree of success. This may be so because the staff-to-student ratio is too high, because the resources are too little, because the procedures used are inadequate or some combination of all three. At any rate before such a project as PIANO is recycled, serious consideration should be given to the adequacy of the program in terms of its objectives and its target population. It is possible that some cherished hypothesis, e.g. that concerning the effect of a devoted and concerned counselor on an underachieving student, needs to be dropped or seriously modified. It is possible that some new thinking might prove helpful. In this regard there is one final recommendation to be made.

In almost every case where an analysis of the difference between schools was made, that analysis indicated statistically significant differences. Unfortunately since the data in hand do not lend themselves to casual inferences, very little can be said regarding the reasons for these differences. And yet these differences may be the most interesting aspect of the data analysis. If some schools were operating programs which were significantly

different in terms of educational procedures and if these different procedures resulted in achievement which was significantly different statistically, then this is precisely the information for which experimental research and development projects are looking.

Unfortunately, neither in the design nor spirit of the PIANO project did inter-school comparisons play a significant role. Perhaps it is not too late to amend this situation. As a final recommendation then it is suggested that during the early fall of 1973, this evaluation team have some block of time (e.g. a week) to present the data it collected and its analysis of *inter-school differences* to the PIANO coordinators and counselors from all six schools. Perhaps together these people might be able to go back over the previous year's procedures, extenuating circumstances and the like and arrive at some educated guesses regarding the causes of the inter-school differences. This effort would be undertaken more in the spirit of research than evaluation and the guesses arrived at could be considered as hypotheses or possible guidelines to be investigated by future personnel.

ANALYSIS OF OBJECTIVES

Objective One

The first objective, as stated in the original evaluation design, was:

> To determine whether at least 75% of the ninth grade students in each school in the program will improve by at least eight school months in reading and arithmetic achievement from April, 1972, to April, 1973, on an appropriate standardized achievement test.

Sampling scheme. The original design stated that achievement scores be collected and analyzed for all ninth grade PIANO students in each school. It was not possible to fulfill this specification. When two sets of tests scores are being compared, the comparison should be done only on students who participated in both test administrations. The range of number of students among the six schools who completed a particular initial and final achievement test varied from a high of 157 to a low of 37 which is approximately 63% and 15% of the students in the program at those respective corresponding schools. The sample sizes of available students were modest so no additional sampling was done. Therefore, the sample consisted of all students at each school whose records showed that they had taken a particular test in 1972, and a second one in 1973.

Sampling bias. The low percentage of participation in achievement test administrations suggests the operation of a self-selection process in taking tests. If so, the scores collected may not be representative of the population of ninth grade PIANO students. Perhaps the more conscientious, more highly motivated, more successful test takers are over represented in this sample. Unfortunately, no evidence is available to resolve this question. All that can be done is to alert the reader to the possibility of such bias.

Test administration and instrumentation. The reliability of achievement test scores for these students, particularly the 1972 scores administered in the junior high schools, was nearly universally questioned by the PIANO personnel in the schools with whom we discussed achievement test scores. Because of this problem several of the schools had readministered achievement tests in the Fall of 1972. These later scores were supplied to us by the schools and were used in computing achievement gains. This resulted in time differences for the 1972 administration. Two schools (B and C) supplied us with reading scores collected in April, 1972; Schools E and F administered reading tests in October, 1972, and Schools A and D did not administer tests until November, 1972. Since more complete and (sup-

posedly) more reliable data were available from these later tests they were used in making the comparison to be reported.

Little hard evidence is available on the question of achievement test score reliability. One indication of unreliability comes from examining the change scores between the 1972 and 1973 administrations of the reading achievement tests. Of 570 students from the six PIANO schools, the change scores of 62 (11%) indicated a *decrease* of more than eight months over the school year. We are not questioning the reliability of this achievement test under certain conditions and for certain populations of students. However, the reliability of the data represented here is questionable.

School A preferred a different type of reading achievement test to the Metropolitan Achievement Test which was the instrument used in all the other schools. They administered the Gates-MacGinitie Reading Survey in November, 1972, and April, 1973. The scores for this test are grade equivalent scores (like the MAT). The correlation on 82 students between the 1973 scores on Gates-MacGinitie and the 1973 MAT was .703. The means were 5.86 and 5.33 and the standard deviations were 1.59 and 1.47, respectively. The distribution of Gates-MacGinitie scores were very similar to the MAT, they were highly correlated, and the Gates scores were more complete and more available. They were therefore used as the reading achievement scores for this school.

The MAT in mathematics was apparently not administered on a city-wide basis in the Spring of 1973 but was to be administered by the PIANO schools to the PIANO students. Only three of the six schools administered mathematics tests. One of these was not comparable to the MAT in either content or in form of the scores, and was therefore not included. Both of the remaining schools administered the MAT test in Fall, 1972 and again in Spring, 1973.

Differences among schools. This evaluation was not charged with comparing the six schools in the PIANO program. Because we suspected that differences among schools would contribute significantly to the error variance, we designed many of the analyses to identify and remove this component of variability. The purpose in so doing was primarily to increase the precision of the analysis. In all cases, the differences among schools were statistically significant, supporting our expectation.

Because a comparison of schools was not requested we did not intend to collect information to differentiate among schools. However, the observations we made at the schools, the general descriptions offered by the PIANO staff, and the data collected all made us realize that the populations of students attending the six schools differed and the programs implemented at the schools differed.

Without identifying particular schools it may be of interest to point out some of the differences observed among the schools. One reason for incorporating this information in the report is that we believe the differences raised questions about the suitability of using the same criteria for evaluating all schools. This point will be elaborated in the section on recommendations.

The reading data (1972 scores) provided one basis for noting that the student populations differed. A one-way analysis of variance on 1972 reading scores indicates that scores do differ significantly among schools. The F-ratio for the effect of schools was 29.8 with 5 and 564 degrees of freedom. The tabled value of F ($\alpha = .05$) is approximately 2.25. Differences among the schools accounted for 21% of the total reading score variance. The mean 1972 reading scores for each school are presented in Table 1.

Analysis of reading achievement data. The stated objective for reading achievement was that 75% of students gain at least eight months in reading during the ten month school year. Because the 1973 reading achievement tests were administered in April, students were in the PIANO program only eight months when this final measurement was taken. The objective therefore requires that 75% of students gain at least approximately the normal rate for this test: one month for each month in school. In the light of these students' previous reading histories is this a reasonable objective?

Table 1 presents information on reading scores for the six schools in the PIANO program. For example, the average 1972 reading score for a student enrolled in School A was 4.94. This average increased to 5.78 in 1973, resulting in a gain of .83 years or 8.3 months. The standard deviations of these distributions are also included in the table. Finaly the percentage of students in each school gaining at least eight months is given. The range of percentages is 31.6 to 46.9. None of the schools are near the objective of 75%.

Analysis of adjusted data. It was previously stated that some of the 1972 tests were administered as late as November. The gains reported for these schools were computed over an interval of only five months. This does not permit a reasonable comparison among these schools.

TABLE 1

Mean and Standard Deviations on 1972 and 1973 Standardized Reading Achievement Tests, Gain Scores, and Percentage of Students Gaining Eight Months or more between the Test Administrations

	Schools					
	A	B	C	D	E	F
1972 Test						
Mean	4.94	7.35	7.04	7.32	5.83	5.20
Standard Deviation	.96	2.15	2.32	2.25	1.73	1.33
1973 Test						
Mean	5.78	8.07	7.48	7.50	6.27	5.70
Standard Deviation	1.53	2.29	2.50	2.24	1.80	1.87
Gain Score	.84	.72	.44	.18	.44	.50
Percentage of Students gaining eight months or more	44.3	41.4	34.1	31.6	46.9	44.3

An estimate of gain for a standard interval of eight months can be computed by dividing the actual gain score of each student by the number of months between the respective test administrations and multiplying this ratio by eight. There is a question whether the most appropriate interval should be eight months or ten months. As the evaluation design was set up (to utilize test administrations already scheduled and regularized for the city), the measure of gain is confounded by combining two months of pre-PIANO education and eight months of PIANO education. If the expectation is that rate of growth in reading will be significantly higher in PIANO than it was prior to PIANO then it would be more straight-forward, simpler, and more conservative to consider only the eight months in PIANO. Therefore an eight month interval was used in computing adjusted scores.

School A had the largest actual average gain of all the schools. These scores were adjusted to see if adjusted scores for the school with the largest gain would meet the reading objectives. The mean increased from .84 to 1.11 meaning that approximately 50% of the students exceeded a gain of eleven months. If 50% exceed a gain of eleven months, did 75% exceed a gain of eight months? The actual gains scores for students in School A were

multiplied by 8/6 and the percentage of adjusted scores equaling or exceeding eight months was computed. The number of scores equaling eight months or more increased from 35 to 48 increasing the percentage who met the objective from 44.3 to 60.8. The adjusted scores for School A still fall considerably short of meeting the stated objective of 75%. Since School A had the largest gain none of the schools would meet the criteria even if adjusted scores were computed.

TABLE 2

Summary of Information on Mathematics Achievement Test Data

	Schools	
	C	E
Sample Size	66	37
Testing Dates	9/72, 5/73	11/72, 3/73
Interval between Tests in Months	8	6
1972 Test		
Mean	5.80	4.96
1973 Test		
Mean	6.69	5.02
Percentage of Students exceeding eight month gain	56.1	18.9
Predicted 1973 scores	6.34	5.25
t-test between predicted and actual 1973 scores	3.87	-1.94

Analysis of mathematics achievement data. As previously mentioned mathematics achievement data was available from only two schools. Pertinent information about these data are presented in Table 2. The gains for neither school is near the objective for mathematics which is the same as that for reading--75 percent of students gaining more than eight months.

Historical regression analysis of achievement in reading. Another method of analyzing gain scores is called the "historical" regression method. This technique uses the 1972 reading score as a measure of reading ability that is assumed to have increased at a constant rate over a student's entire length of time in school. Using this assumption, an average reading gain per school month is computed for each student. This average gain is then multiplied by the number of months in the PIANO program prior to the 1973 test. In this way an estimate of achievement gain for the length of time in PIANO is computed based on a student's previous gain in (supposedly) conventional programs. If the PIANO program is effective the actual gains achieved in this program should be higher than that extrapolated from previous experience in conventional programs. A correlated t-test is used to compare predicted versus actual 1973 scores. A statistically significant difference (in favor of the actual score) is accepted as evidence that the PIANO program was an effective one.

The data for performing an historical regression analysis were computed and are displayed in Table 3. The 1973 actual score is higher than the predicted score for all schools except Schools C and D. However the amount of the difference must be sufficiently large to conclude that there is a real difference between test score distributions for the population of PIANO students. If the computed value of the t statistics is greater than the tabled value of t for the appropriate degrees of freedom and for the level of significance specified, then we can conclude that the differences between the populations are large enough to be considered real ones. The tabled value for t (one-tail test) with level of significance of .05 and 80 degrees of freedom is 1.665.

Students in School A increased in reading ability an amount greater than that predicted by the historical regression technique. School F falls just a hair below the cut-off value for t so it will be considered to have also increased by a significant amount.

Although it was not anticipated that a school might fail to gain as much as predicted, such was the case for School D (for a two-tailed test with $\alpha = .05$, and 100 degrees of freedom the tabled value of t is 1.984). Students in the remaining three schools did about as well as predicted by the historical regression method.

In case there is any doubt in the reader's mind the 1973 reading scores are higher than the 1972 scores. Even School D produced a t of 1.696 on a correlated t-test. The tabled t for 100 degrees of freedom and $\alpha = .05$ is 1.660 (one-tail test).

TABLE 3

Summary of Information on Historical Regression Method of Analyzing Reading Achievement Data

	Schools					
	A	B	C	D	E	F
1973 Actual Score	5.78	8.07	7.48	7.50	6.27	5.70
1973 Predicted Score	5.23	8.00	7.66	7.70	6.18	5.46
Actual Gain	.84	.72	.44	.18	.44	.50
Predicted Gain	.29	.65	.62	.38	.35	.26
t-values	5.834	.697	-1.330	-1.816	.418	1.66
Sample Size	79	157	91	114	49	80

Historical regression analysis of achievement in mathematics. The same procedure described above was applied to the mathematics data. The results are reproduced in Table 2. The computed value of t for School C (3.87) is larger than the appropriate tabled value (1.671). The average gain of 8.9 months is significantly higher than the predicted gain of 5.4 months. Failure to gain the predicted amount in School E is almost significant.

Discussion of historical regression technique. The historical regression method is superior to a fixed criterion in at least one way. This method is sensitive to individual past history. The criterion on which a student is judged is one established by his own past performance. If a student, during the experimental period, does better than a measure predicted by his past history, the experimental treatment is judged to be effective.

The greatest problem with this method is to obtain a reliable estimate of past achievement gains. The use of one score, and one that is suspected as being unreliable, to characterize the average growth in achievement over a long time span is not a satisfactory procedure. More data points are needed to develop a fair and reliable estimate of growth.

The assumption of a linear growth rate for students such as these in the PIANO program might also be questioned. An equally likely hypothesis might be that these students exhibit a negatively accelerating growth curve. An investigation of growth curves for a sample of students from the population over a period of time is needed to validate the shape of the growth curve appropriate for use in studies like this one.

Summary. Only one school (A) approached the stated objective for reading achievement, when gains were adjusted to an eight month interval, but did not meet it. Two schools (A and F) exhibited reading gains greater than that predicted by the historical regression technique. Of the two schools that provided mathematical achievement scores, neither achieved the stated objective but one school (C) achieved higher scores than that predicted by historical regression. The schools that did achieve higher scores in mathematics and reading are not the same schools. The different results achieved by analyses raise questions about the setting of criteria for evaluation that will be discussed in the recommendations section.

It is interesting to point out that the school (A) with the lowest entering scores of all six schools had the largest actual gain in reading of all schools. If the PIANO program were the same at all schools this finding would be very unexpected. This result suggests that something different was happening at School A. From what we know of the reading programs, the one at School A was a more structured, remedial program than was true of programs at other schools. Perhaps the goals of the reading program at School A were more nearly matched to the reading achievement tests than were the programs at other schools.

For both the reading and mathematics areas, one school in each area, and not the same school, failed to maintain the previous growth rates of students in these respective areas. This finding further supports the previously expressed conviction that the programs at the PIANO schools are not the same programs. Much more might be learned about educational efforts similar to the PIANO program by comparing and analyzing the differences implemented by different schools in the program and the effects associated with these differences by ignoring these differences.

Objective Two

The original objective was:

> To determine whether at least 75% of the ninth grade students in the program in each school will improve by at least two scale points in appropriate school behavior from initial to final rating by the grade advisor.

The design proposed that a ten item, five-point rating scale be developed by the evaluation agency and that it be used by the grade advisor to evaluate all PIANO students in October and again in May. "Appropriate school behavior" was not defined by the design. Examples of items suggested in the design were:

1. Whispers or talks in class
2. Talks back to teacher
3. Withdrawn from groups

Since the original design could not be followed, because of the lateness in contracting for the evaluation, the pre-post test design was changed to one of comparing PIANO students with a group of non-PIANO students who were as similar to PIANO students as possible. Attending to or participating in classroom activities was judged to be the essential aspect of student behavior sought by the original design. Therefore it was selected as the behavior to be measured via a modified Jackson-Hudgins Observation Schedule (Labadene, H., 1967) with a PIANO and non-PIANO group in each school.

Procedure for using Jackson-Hudgins Observation Schedule. The procedure was for the observer to "sweep" the room, looking at each student in turn, and immediately recording the students' state of attention on a matrix observation schedule prepared in advance. Four classifications were possible: a plus was recorded if the student was attentive; a minus was recorded if the student was clearly inattentive; a question mark indicated that the observer was uncertain what the student's state of attention was; and a zero signified that attention was not observable. The observer continued to sweep the students in the classroom for an entire class period. The number of completed sweeps varied from 15 to 50 depending on the number of students in the classroom. The sum of the tallies entered for each of the four attention categories were totaled for each student in the classroom to obtain an attention profile.

Inter-observer reliability, defined as percentages of agreement between two or more

judges, ranged from 86 percent to 99 percent for a series of observations made by the originators of this observation schedule.

Classroom observation. At least two classrooms were observed in each of the six schools. The PIANO staff selected one PIANO class and attempted to match it with an equivalent non-PIANO class. The observations were completed within a week and a half to minimize the effect of uncontrolled influence on behavior.

Analysis of data. Since the number of question marks and zeroes was quite small they were dropped from the analysis and attention was focused on the frequencies of pluses and minuses or states of attention and inattention, respectively. As can be seen in Table 4 the overall behavior of the non-PIANO students was significantly different and in this case significantly more attentive than the PIANO students. The differences in frequencies of states of attention and inattention amongst these two groups of students with an N = 187 yielded a chi-square of 65.5, well beyond chance expectations.

TABLE 4

Chi-square Analysis of Attentiveness of PIANO and non-PIANO Students

	Groups	
	PIANO	Non-PIANO
Attentive	1601	1914
Non-Attentive	573	377

However, an examination of the results, school by school, indicates that in three schools the PIANO students were significantly more attentive than non-PIANO students while in only two schools was the reverse situation found. This school-by-school look also indicates that when PIANO students were inattentive they were *very* inattentive, to a much greater degree than the inattentive non-PIANO students. This extreme difference in degree in the two schools accounts for the overall difference between PIANO and non-PIANO students. A further analysis of the conditions in School A which differed in the extreme and accounted for the greatest part of the overall difference indicates an unusual condition. Specifically, the room temperature at the time of the observations was above 95°, a

condition which would stimulate inattention on the part of most high school students. However, the non-PIANO group was highly motivated that day due to the fact they received and discussed their final examination of the term. Allowing for this atypical set of circumstances and dropping School A from the analysis, PIANO students were not significantly more inattentive than non-PIANO students. The chi-square value was .097. To be significant chi-square must exceed 3.84 with one degree of freedom and $\alpha = .05$. However, the weakness in these procedures is apparent. This observation schedule is sensitive to the nature of the lesson being observed as well as to the style of the teacher and any atypical conditions which might prevail. Therefore, in order for valid conclusions to be inferred from data collected with this schedule it should be used with a much larger, more representative sample of classroom lessons for both PIANO and non-PIANO groups.

Objective Three

The original objective was:

> To determine whether at least 75% of the ninth grade students in the program in each school will improve by at least two scale points in motivation and positive attitude toward school from initial to final rating by the grade advisor.

The original design proposed that the evaluation agency prepare a ten item, five-point rating scale. The grade advisor was to rate every ninth grade PIANO student in the school in October and again in May on "motivation and positive attitude toward school." Four sample items suggested in the design were:

1. Having self confidence
2. Having good work habits and study habits
3. Wanting to learn and improve himself
4. Volunteering to do things in class

The proposed five-point multiple choice response ranged from "Always" to "Never." The original pre-test post-test design could not be implemented. A one-time comparison between PIANO and non-PIANO students was proposed as an alternative.

Attitude toward school. Two different attitude questionnaires were selected to be administered to samples of PIANO and non-PIANO students in each of the six schools.

The Student Opinion Poll (SOP) developed by Getzel and Jackson (1962) was selected to measure attitude toward school. The forty nine items on the test measure attitude toward teachers, the curriculum, the student body and classroom procedures. The student checks the one response of three or four alternative responses to each question that best matches his opinion of the school he is attending. One response is designated as the most satisfactory response. Each time this response is selected a score of one is given.

A sample item from the questionnaire is:

13. Most of the subjects taught in this school are
    a. very interesting
    b. above average in interest
    c. below average in interest
    d. dull and uninteresting

A copy of the entire questionnaire is included in Appendix A. The questionnaire was originally prepared for use with sixth graders. One reason for selecting this instrument was the expectation that this questionnaire would be readable by students in the PIANO program. Reliability of the instrument from previous administrations with sixth grade students varied from .85 to .89.

One change was made in the questionnaire. An item that read "as preparation for high school, the program of this school is . . . . " was changed to replace high school with college.

Objective for attitude toward school. It was very difficult, with the amount of information available to us about both classes of PIANO students and classes of non-PIANO students that would be judged to be similar to PIANO classes students, to establish expectations about differences in attitudes toward school between these two groups. No directional hypothesis was therefore stated for attitudes toward school. Instead data was sought to determine whether the attitudes toward school as measured by the SOP would be significantly different for PIANO students as compared to non-PIANO students.

The analysis of data should be a t-test or appropriate Analysis of Variance to compare the PIANO and non-PIANO students. The results of the analysis will be reported below. The questionnaire was to be administered in May or June 1973 to selected classes of PIANO and non-PIANO students. It was understood that the non-PIANO classes closely equivalent to PIANO classes do not exist in most of the schools but that classes most similar to the selected PIANO classes would be used as reference groups. A convenient sample of classes, although not desirable, was accepted as necessary to minimize the practical problems in obtaining classes for collecting this information.

Attitude toward counseling services. The staff at several of the schools seemed to feel that one of the most effective parts of the PIANO program was the special counseling services provided for PIANO students. We selected, as a second attitude instrument, a questionnaire designed by Form (1955) to assess attitudes concerning counseling services available at school. The original questionnaire is a twenty two item scale with five response alternatives from "Strongly Agree" to "Strongly Disagree." The student responds by circling the response that indicates how he feels about the services. A sample item is:

It is a complete waste of time to go to the counseling office.

Strongly Agree Agree ? Disagree Strongly Disagree

The split-half reliability of the instrument on a sample of 544 college students was .94. One item was dropped from the instrument and the wording of several items was simplified. Specific names were also changed to use terminology more commonly associated with counseling facilities at high schools. A copy of the questionnaire is included in Appendix B. Each item on the Counseling Attitude Scale was scored from one to five depending on which of the five responses was checked by the respondent. The possible total score range of the instrument was 21 to 105.

Objective for attitude toward counseling service. The objective for counseling attitude was to demonstrate that the PIANO students had more favorable attitudes toward counseling facilities at school than did the non-PIANO students. The method of data analysis was a t-test or an Analysis of Variance. Administration of the questionnaire and sampling procedures were to be the same as specified for attitudes toward school.

Sampling scheme. The PIANO staff at each school was requested to select one PIANO class and one non-PIANO class, that was as similar to the PIANO class as possible, to respond to the questionnaires. In a few cases, when class sizes were small, additional classes were used to increase the sample size. The number of students responding to the questionnaires was 399 on the Student Opinion Poll and 385 on the Counseling Attitude Scale. A schedule for administering questionnaires was arranged with the PIANO staff at each school. One class period was sufficient for students to respond to both questionnaires.

Analysis of attitudes toward schools. Responses to the Student Opinion Poll were collected from 196 PIANO and 203 non-PIANO students. The questionnaires were scored and analyzed with a two-factor analysis of variance. The two factors were groups (PIANO and non-PIANO) and schools. The analysis of variance table is presented in Table 5.

There was no difference in responses between the two groups. The percent of variance accounted for by group membership was very small (.1%). There was a significant difference among schools. Schools accounted for 38% of the total variance. The mean attitude scores for each school, in the order School A to F, are: 19.9, 20.2, 16.3, 19.1, 20.0, 18.2.

Analysis of attitudes toward counseling services. The Counseling Attitude Scale was responded to by 193 PIANO and 192 non-PIANO students. The same analysis procedure was applied to these data as to the Student Opinion Poll. The analysis of variance table for these data is presented in Table 6.

TABLE 5

Analysis of Variance for Student Opion Poll

	Sum of Squares	Degrees of Freedom	Mean Square	F ratio
Groups (PIANO – Non-PIANO)	10.325	1	10.325	<1
Schools	776.354	5	155.271	3.31
Error	18.405.392	392	46.95	
Total	19.181.746	398		

Tabled F: (α = .05, df = 5,120) = 2.29

TABLE 6

Analysis of Variance for Counseling Attitude Scale

Effect	Sum of Squares	Degrees of Freedom	Mean Square	F ratio
Groups (PIANO – Non-PIANO)	507.871	1	507.871	3.1
Schools	917.448	5	183.489	1.1
Error	61,407.216	378	162	
Total	62,832.535	384		

Tabled F: (α = .05, df - 1,120) = 3.92

The effect of neither of the two factors analyzed were large enough to reject the null hypothesis of no difference among levels of the factors. It is concluded that there were no differences in attitudes toward counseling services for either of the two groups or among the six schools. The objective stated for attitudes about counseling services was not attained.

Summary. PIANO and non-PIANO students did not differ in attitude toward school in general or in attitude toward the counseling services. The expected higher attitude toward counseling service by PIANO students was not supported by the data. There were differences toward school among the schools but such differences were not found in attitude toward counseling service.

Objective Four

The original evaluation design proposed:

> To determine whether there was a 30% increase in attendance of the ninth grade students in the PIANO program in each school. The sample was to be the entire population. Attendance was to have been measured and compared between an initial period of October and November 1972 and a final period of April and May 1973. A percentage comparison between the two periods was requested.

Sampling scheme. It was proposed, and accepted, that rather than sample attendance at an initial and final period, samples would be taken throughout the school year to determine if attendance patterns were subject to seasonal variations. For example, attendance may increase gradually over the year but drop suddenly in late April or May because of weather conditions. The sampling scheme originally proposed would not detect such variations. By monitoring one week randomly selected from each month these kinds of patterns can be identified and the amount of variation determined.

It was also proposed, and accepted, that data not be collected for the entire population but that a random sample of students be selected; that data be collected on this sample, and that these data be analyzed to test this objective. We decided to eliminate the first (September) and last (June) months of the school year from the population because they may contain unusual effects associated with the opening and closing of school. January was also dropped because of the large number of days in the month when school was closed. For the remaining seven months, all weeks not including holidays were identified and one week each month was randomly selected as the sampling period for the month. The periods selected were: October (2-6), November (13-17), December (4-8), February (5-9), March (26-30), April (2-6), and May (14-18).

A sample of 25 male and 25 female students were selected from each ninth grade PIANO school population. The population was defined as students who were in the PIANO program throughout the year and who were freshmen for the first time in September 1972. The equal sampling of male and female students was done in order to look for comparisons between sexes. The list of students who comprised the population was prepared and nth student was selected: n being an appropriate number to produce approximately 25 students from one pass through the list. In most schools additional names had to be sampled from

the list because records of one kind or another could not be found or a selected student had recently moved or been transferred from the program.

Analysis of attendance data. The attendance data was recorded and analyzed as number of absences during one five-day week. Table 7 displays the average number of absences per sampling period over the year by sex and school. For example, the average number of absences per sampling period over all seven sampling periods for males at School C was .73 days. Females at this school had a nearly equal rate of absenteeism, .71 days per sampling period.

TABLE 7

Average Number of Absences by Sex and School

		Schools						Average for All Schools
		A	B	C	D	E	F	
Sex	Male	1.71	.78	.73	.55	.99	1.81	1.00
	Female	.94	.96	.71	.44	.93	1.26	.87

A repeated measures four way analysis of variance was computed on the absence data. The four factors were the seven sampling periods, six schools, two sexes and students. (Students appear as factors in repeated measures designs. The variation among students is nearly always significant if it is evaluated. The effect of this factor will be ignored in this analysis and also in those in the following objective.) Repeated measures were over the levels of sampling periods. Details of the analysis, which is somewhat complex, are described in Appendix C. The result of the analysis demonstrates differences among sampling periods, among the schools, and a sex by school interaction effect. The objective of this analysis was to determine whether attendance increased during the year. The variations among sampling periods indicate how attendance (or rather absences) changed over the year.

There was a statistically significant difference among the sampling periods. The mean number of absences per student per sampling period, beginning with October and ending with April, was 0.71, 1.09, 1.20, 0.86, 0.95, and 1.20. It appears that absences start out low each semester and increase as the semester progresses. The same pattern occurs in both the Fall and Spring semesters. (The reader should recall that January was not sampled so the 0.86 datum is for February.)

Since the design called for it, a comparison was made between the October-November average attendance and the April-May average attendance. The absence data was converted to mean number of presences per student per sampling period by subtracting absences from total number of days in the sampling period (5). The average number of presences was 4.29, 3.91, 3.80, 4.14, 4.12, 4.05, 3.80. The average attendance during October-November is 4.10 and during April-May is 3.925. The change is a decrease, not an increase. The percentage is 4.27. Although a statistical test was not supplied to this comparison this difference is so small it is not likely that the differnce is significant and it should not be interpreted as a meaningful change in attendance. The appropriate conclusion is that attendance was no different at the end of the year than it was at the beginning. The goal of the objective was, of course, not met.

The analysis showed that there were differences in number of absences among the six schools. In order, School A through School F, the average number of absences per sampling period over the year were 1.33, .87, .72, .50, .96, 1.53. School F had three times as many absences per sampling period as did School D.

Although the difference between sexes was not quite large enough to be significant at the .05 level of confidence, the sex by school interaction was significant. The data of this comparison is reproduced in Table 7. In Schools A and F the boys were absent more than the girls. In Schools C, D and E the number of absences for both sexes was very similar. In School B the girls were absent more than the boys.

Summary. There was a change in attendance among the sampling periods throughout the year. The change was not of the nature stated as the objective for these data. It did not increase from the beginning to the end of the year. Attendance appears to have been highest at the beginning of each semester of the year and to have decreased as the semester progressed. The goal of this objective was not met.

There were differences in attendance among the six schools and there was a significant interaction between schools and sex. In two schools (A and F) boys had more absences than girls. In School B girls were absent more than boys; in the remaining schools the differences between sexes were small.

Objective Five

The original evaluation design established a goal of a decrease of at least 35 percent in disciplinary infractions of ninth grade students in the program over the school year. All ninth grade PIANO students in each school were to be used. The number and type of disciplinary referrals for the period October and November 1972 were to be compared with April and May 1973. A percentage between periods was to be done.

This objective was changed to sample the student population and to expand the sampling periods of disciplinary data from two two-month periods to four two-month periods: October-November, December-January, February-March, and April-May. The same sample of students used in Objective Four was used for this objective.

In all schools, except one, disciplinary matters were the province of the deans. In one school disciplinary infractions of PIANO students were referred to the PIANO counselor. There were wide variations from school to school in the type and extent of records kept. The only classification scheme that developed out of discussions with several of the school personnel was to identify as major infractions referrals that resulted in suspension and to classify all others as minor infractions. The records for each student in the sample were located in the Dean's files. Every referral was examined and classified as major or minor and entered in the appropriate sampling period according to the data of the referral.

Analysis of disciplinary infractions. The number of major infractions was very small. A total of seven major infractions were reported among the six schools. Six of these were reported by School E and involved four infractions by girls (two by the same girl) and two by boys. Four of the seven infractions occurred in the April-May reporting period. There was not a sufficient number of major infractions to do any further analysis of these data.

A repeated measures analysis of variance (similar to the one used in Objective Four) was performed on the minor infraction data. The four factors of the analysis were the four sampling periods (with repeated measures over this factor), the six schools, sex, and students. The detailed strucutre of the analysis is presented in Appendix D. The result of the analysis was to demonstrate differences among the sampling periods, among the schools, between sexes, and an interaction between sampling periods and schools.

The average number of infractions changed significantly over the four sampling periods from 0.82 to 0.62 to 0.91 and to 0.89. The December-January sampling period was much

lower than the other three. However, there was not a drop from the first period to the last period. An increase of 8.5% was computed between the beginning and ending sampling periods but the difference is so small it is not likely that it is significant. With the exception of the December-January period the number of infractions remained fairly constant. In any case the objective of a 35% decrease was not achieved.

There were other differences in the data. The number of infractions among the schools differed significantly. For schools A to F the average number of infractions per student per sampling period were .34, 1.20, 1.48, .52 .43 and .90. More than four times as many infractions were reported at School C than at School A. It may be of interest to note that School A, where discipline for PIANO students was administered by the PIANO counselor, reported the lowest number of infractions. Not enough is known about the disciplinary procedures from school to school to relate these differences to specific practices.

There was a difference between sexes in the number of minor infractions. The average number reported for boys over the year was .97 and for girls .66.

There was also a significant interaction between schools and sampling periods. Such a finding indicates that the variation in minor infractions across sampling periods is not the same for each school. While infractions increased over the year for some schools, it decreased for others and was nearly constant for still others. These data are also difficult to interpret without more detailed information about each school's disciplinary situation.

Summary. The goal of objective five was not met. The only change in the average number of minor disciplinary infractions over the year was a decrease for the December-January sampling period. Males were recorded as having committed more minor infractions than females. Significant differences among schools and an interaction effect between schools and sampling periods were not interpreted for lack of information about disciplinary practices at each school. The number of major infractions (suspensions) were too few to analyze.

Objective Six

The original evaluation objective stated that there would be an increase in home and parental contacts leading to a resolution of student referral problems.

By the time this evaluation began, the bulk of the family-home contacts had already been completed. Unfortunately, the evaluators did not have any opportunity to influence the keeping of records relating to these contacts. Consequently, although in most cases there appears to have been a very large and useful amount of energy expended by the PIANO staff, in particular by the Family Assistance Counselors, the records of these activities are different from school to school, generally informal or anecdotal and on the whole extremely difficult to interpret in terms of the orginally stated objective. For example, in many cases it was not possible to make a determination from the available records whether a problem was successfully resolved or not. To make this interpretation would have required lengthy consultations with the Family Assistance Counselors at each school. In most cases either she or the evaluation staff felt that time spent in such pursuit would be too costly in terms of responsibility to other phases of the project. As a result we can only report that many discussions with PIANO staff, school administrators, Family Assistance personnel as well as examination of available records, all indicate that these activities were considerable in effort and achievement. On this basis there seems little doubt that such activities should be continued but that consistently uniform and specific record keeping should be encouraged. In contrast, we might note in passing, the general conversations revealed no such enthusiasm concerning the efforts and achievements of some of the school psychologists connected with the project. Nor, for that matter, was a specific objective related to this project explicitly designated for evaluation.

Objective Seven

The orginal design stated in its objective for class grades that 60% of all ninth grade PIANO students in each school pass all their subjects in January 1973 and that this increase to 75% passing all subjects in June 1973.

The only change in this objective was to sample the student population rather than to collect data for the entire population. The sample of 50 students from each school on which attendance and disciplinary data was collected was used to collect grades.

The raw data of this analysis was the proportion of courses passed. This form was used for two reasons. It was relatively easy to classify recorded marks as pass or fail but it was difficult to arrive at numerical scores for each course. Secondly, this form was convenient to use for reporting percentages passing all classes.

Analysis of grades. A similar four factor repeated measures analysis of variance was computed on the grades data. A detailed description of the analysis is available in Appendix E. The result of the analysis is that there were significant effects due to sampling periods (January and June), schools, sex, and a sex by sampling period interaction. The sex by school interaction was not evaluated because an appropriate denominator term was unavailable for the analysis.

There was a difference in grades between the two sampling periods. The average proportion of courses passed decreased from January to June. The mean January proportion was .66; the June mean was .57. About 50% of the students passed two-thirds of their classes in January but only three-fifths of them in June. Table 8 shows the percentage who passed all courses in January and June at each school. Obviously the goal of this objective was not met.

TABLE 8

Percentage of students in each school who passed all courses in January and June

	A	B	C	D	E	F	Total
January, 1973	2	36	47	46	36	22	31.3
June, 1973	2	23	42	42	26	4	24.0

The other results of the analysis indicates that the proportion of courses passed varied significantly among schools. In the order School A to School F the average proportions passed for the year were .35, .74, .75, .77, .53, .55. Approximately 50% of the students in each of these schools passed at least the proportion of classes listed above. These percentages are based on the assumption that the mean and median do not differ substantially. The assumption was not tested.

The difference in grades between sexes was in favor of females who passed an average of .66 courses while males passed .57 of their courses. Finaly there was a sex by marking period interaction. The data of this interaction is presented in Table 9. Although both sexes decreased in proportion of courses passed from January to June the size of the decrease was much larger for males than for females.

TABLE 9

Average proportion of courses passed by Sex and Marking Period

	Males	Females
January	.64	.69
June	.50	.63

Summary. There was a change in grades between January and June but it was a decrease rather than the expected increase. Even in January the percentage of students who passed all their courses was 31.3 instead of the expected 60. Additional differences were noted between sexes, with females passing more courses than males. The change in the proportion of courses passed from January to June differed for sexes. The males decreased more than the females did. There were also differences in grades among schools. Students in Schools B, C and D, on the average, passed twice as many courses as students in School A. It is interesting to note in passing that School A is the school that had the greatest increase in reading of all schools.

RELATIONSHIPS AMONG ATTENDANCE, GRADES, AND DISCIPLINARY INFRACTIONS

By collecting attendance, grades, and number of disciplinary infractions on each of the 50 students sampled from each school, we are able to indicate how these three variables related to each other. Combining all schools and sexes, the correlation over the total sample was:

Correlation between Grades and Absences	−.652
Correlation between Grades and Disciplinary Infractions	−.217
Correlation between Absences and Disciplinary Infractions	−.103

For the size of this sample (570) a correlation must exceed ± .09 to be significant at the .05 level of confidence. Thus all three of these correlations are significantly larger than zero, although the relation between grades and number of absences is much larger than the others.

A high negative correlation between grades and absences indicates that a student with a large number of absences is likely to have passed a low proportion of courses. Stated in another way, a student who was present most of the time is likely to have passed a high proportion of courses. No causal relation should be inferred from these data. High attendance is not to be interpreted as the course of higher grades or vice-versa. A modest negative correlation such as that between grades and disciplinary infractions also indicates a pattern of higher number of infractions being paired with lower proportion of courses passed. However the consistency of this relation is not as strong as with a high correlation. The probability of the relation holding is less than is the case for a high correlation.

Computing these correlations separately by sex shows very little difference between absences and grades (females, −.65, males, −.66) but sexes vary in the relation between grades and number of disciplinary infractions. The correlation is higher for females (−.262) than for males (−.170). There is no correlation between absences and number of disciplinary infractions for males (.022) although for females there is a small positive correlation (.239). The correlations by school and sex are reproduced in Table 10. Within a given set of variables, the correlations for females vary more among schools than is true for males. As before, it is difficult to interpret these differences without more knowledge about inter-school differences.

TABLE 10

Correlations among Absences, Disciplinary Infractions, and Course Grades by Sex and School.

School	Absences and Course Grades		Absences and Disciplinary Infractions		Course Grades and Absences	
	Males	Females	Males	Females	Males	Females
B	−.697	−.895	.249	.427	−.502	−.614
D	−.531	−.799	.123	.308	−.299	−.435
C	−.745	−.755	.123	.590	−.544	−.613
A	−.574	−.659	−.099	.025	−.298	−.216
E	−.620	−.600	−.029	−.098	−.010	−.236
F	−.763	−.506	.010	.240	−.394	−.331

RECOMMENDATIONS

The recommendations in this section fall into two general categories, those concerned with evaluations of projects of this type, and those concerned with the design of future projects.

Concerning evaluations

1. It is strongly recommended that the personnel of new programs that are to be evaluated should be directly and seriously consulted in the development of evaluation objectives, criteria, and measurement procedures. Not only might this result in better staff morale and attitude toward evaluation but it might also result in better evaluation designs. For example, had the project staff been consulted, they would have emphatically pointed out that in the original evaluation design it was completely unrealistic to base the evaluation of the changes in attitude and behavior of 250 students on the judgment of a single grade advisor. It is unlikely that a single person could have an adequate enough knowledge of such a large group of students to be able to rate them accurately and reliably.

2. Objectives should be supported by some rationale or hypotheses and should be based on past data, when such data are available. For example, the expectation of the achievement of 8 months growth in as many months as measured by a standardized test normed on a national sample, may be unrealistic for a group of learners that has been gaining at a consistently slower rate over a period of many years. As previously discussed, objectives based on a historical regression approach may be more sensible in this situation.

Similarly, there probably exist extensive and relatively reliable attendance data for the students in experimental projects from which future trends, in terms of such factors as schools, seasons, and sex, could be predicted with relative accuracy. Changes in these trends due to some programmatic treatment could then be equated with "effect size" in experimental research. Hypotheses regarding the amount of change expected could then be offered in terms of the standard deviation of past distributions of attendance data. For example, a large change in attendance could be an increase of a single standard deviation over last year's average for some group. A small change might

be a quarter of a standard deviation and so on. The point being that the objective would be stated in terms related to past data, current recommended research procedures, and based on some hypotheses or rationale for expecting an effect of a specific size.

3. Objectives should be sensitive to differences in goals and procedures at different project units. As discussed previously, different schools in the PIANO project operated fairly divergent programs and yet all were evaluated by the same criteria. Furthermore, it is recommended that where there are differences among units, some systematic attempt be made to analyze the possible effects of these differences.

4. The persons who will actually implement the project evaluation should have the opportunity to interact with the project personnel at the very early stages of the project. This early involvement may be important for the following reasons:

a. there would be agreement on the precision of the data to be collected, times of test administrations, and formats of records for recording and storing data, thus making the data collection procedures more efficient and less costly and allowing the data analysis procedures to be more powerful.

b. the improvement in efficiency and effectiveness of the data collection and analysis procedures could lead to interim or formative judgments which could be fed back into the project in time to guide its procedures on the basis of its current effectiveness or lack of it (e.g. if it was realized that the remedial reading program at School A, or the remedial math program at School C, or the attendance rate at School D was significantly different from those at other schools, the "causes" of these differences might have been inferred and generalized to the programs at the other schools);

c. the formative or guiding effect of the evaluation data might have initiated a feeling among the staff that evaluation procedures were an integral and helpful part of the project rather than merely a summary judgment of how well they had done their jobs;

d. procedures for monitoring and describing the programs at each school, in particular their differences, could have been established to guide each unit in pursuing the uniqueness of its objectives.

APPENDIX A

Student Opinion Poll

STUDENT OPINION POLL

This is not a test. The answer to each question is a matter of opinion. Your true opinion, whatever it is, is the right answer. You will be asked a lot of questions about the school in which you are now studying. Wherever the words "school," "teacher," and "student" appear, they refer to *this* school, the teachers you have had while studying *here*, and your classmates in *this* school.

Here is an example:

Circle the letter to the left of the answer you choose.

0. In general I study
   a) too little
   b) too much
   c) about the right amount

If your answer is "a) too little," circle the 'a' in the first answer, like this:

   (a) too little
   b) too much
   c) about the right amount

If your answer is "b) too much," circle the 'b' in the second answer, like this:

   a) too little
   (b) too much
   c) about the right amount

If you have any questions, raise your hand and you will be helped.

Project for Increased Achievement and a New Outlook
Directors: Donald Payne, Robert Ruderman
New York University
Center for Educational Research and Field Services

1. This school listens to parents' opinions

    a. too much
    b. just enough
    c. too little

2. The number of courses given in this school is

    a. too many
    b. just about right
    c. not enough

3. Although teachers differ in this school, most are

    a. very good
    b. good
    c. fair
    d. poor

4. In some schools the principal sees and talks with the students often, while in other schools he rarely sees them. In this school the principal sees and talks with students

    a. too often
    b. just about the right amount
    c. too little

5. The chance to say or do something in class without being called upon by the teacher is

    a. too little
    b. too much
    c. about right

6. The things that I am asked to study are of

    a. great interest to me
    b. average interest to me
    c. of little interest to me
    d. of no interest to me

7. Getting to know other kids in this school is

    a. easier than usual
    b. about the same as other schools
    c. more difficult than usual

8. As preparation for college, the program of this school is

    a. too tough
    b. about right
    c. too easy

9. The class material from year to year

    a. repeats itself too much; you learn the same material over and over
    b. repeats itself just enough to make you feel what was learned
    c. is so new that the things learned in the last grade do not help much in this one

10. In this school the teachers' interest in the students' school work is

    a. too great
    b. just about right
    c. not great enough

11. When students in this school get bad grades, their classmates usually

    a. feel sorrier for them than they should
    b. admire them more than they should
    c. show the right amount of concern

12. Students in this school are

    a. too smart – it is difficult to keep up with them
    b. just smart enough – we are all about the same
    c. not smart enough -- they are so slow I get bored

13. Most of the subjects taught in this school are

    a. very interesting
    b. above average in interest
    c. below average in interest
    d. dull and uninteresting

14. The teachers' interest in what the students do outside of school is

    a. too great
    b. about right
    c. too small

15. The student who shows a sense of humor in class is usually

    a. admired by the teacher more than he should be
    b. punished by the teacher more than he should be
    c. given about the right amount of attention

16. When teachers "go too fast," students do not know what is going on. In this school, most teachers teach

    a. too slowly
    b. about right
    c. too fast

17. Students who are good in sports are respected by classmates

    a. more than they should be
    b. less than they should be
    c. neither more nor less than they should be

18. The practice of competing against each other or of working together in this school

    a. leans too much toward competition
    b. leans too much toward working together
    c. is well balanced

19. On the whole, the things we study in this school

    a. are about right
    b. should be changed a little
    c. should be completely changed

20. The teachers I have had in this school seem to know their subject matter

    a. very well
    b. quite well
    c. fairly well
    d. not as well as they should

21. Students may work either by themselves or in groups. In this school we work in groups

    a. too often
    b. just enough
    c. too little

22. Students get along in this school

    a. very well
    b. about average
    c. not too well
    d. very badly

23. The amount of "school spirit" at this school is

    a. more than enough
    b. about right
    c. not enough

24. On the whole the school pays attention to the things you learn from books

    a. too much
    b. just enough
    c. not enough

25. Teachers in this school seem to be

    a. almost always fair
    b. generally fair
    c. occasionally unfair
    d. often unfair

26. The things we do in class are planned

    a. so badly that it is hard to get things done
    b. so well that we get things done
    c. so completely that we hardly ever get to do what we want

27. Our seats in class

    a. change too much; we can never be sure where we will sit and who will sit next to us
    b. change about the right number of times
    c. never change; we stay in the same place all year

28. The students who receive good grades are

    a. liked more than they should be by their classmates
    b. disliked more than they should be by their classmates
    c. neither liked nor disliked more than they should be

29. In this school the teachers' interest in the students' school work is

    a. just about right
    b. not great enough
    c. too great

30. In my opinion, student interest in social affairs, such as clubs, scouts, and the "Y" is

    a. too great
    b. about right
    c. too little

31. In general the subjects taught are

    a. too easy
    b. about right in difficulty
    c. too difficult

32. When students need special attention, teachers in this school are

    a. always ready to help
    b. generally ready to help
    c. ready to help if given special notice
    d. ready to help only in extreme cases

33. The ability of the teachers in this school to present new material seems to be

    a. very good
    b. good
    c. average
    d. poor

34. In general, students in this school take their studies

    a. too seriously
    b. not seriously enough
    c. just about right

35. In this school teachers seem to teach

    a. too many things that are *not* useful to us now
    b. too many things that are useful to us now but not later
    c. both things that are useful now and can be useful later

36. When it comes to grading students, teachers in this school are generally

    a. too "tough"
    b. just "tough" enough
    c. not "tough" enough

37. The student who acts differently in this school is likely to find that most students

    a. dislike him for being different
    b. do not care whether or not he is different
    c. like him for being different

38. In my opinion, students in this school pay attention to their looks and clothes

    a. too much
    b. about right
    c. too little

39. In general, teachers in this school are

    a. very friendly
    b. somewhat friendly
    c. somewhat unfriendly
    d. very unfriendly

40. In general, I feel the grades I received in this school were

    a. always what I deserved
    b. generally what I deserved
    c. sometimes what I did *not* deserve
    d. frequently what I did *not* deserve

41. Teaching aids such as films, radio, and the like are used

    a. more than they should be
    b. as much as they should be
    c. less than they should be

42. Memory work and the learning of important facts are

    a. stressed too much
    b. used about right
    c. not stressed enough

43. In some classes the teacher is completely in control and the students have little to say about the way things are run. In other classes the students seem to be boss and the teacher contributes little to the control of the class. In general, teachers in this school seem to take

    a. too much control
    b. about the right amount of control
    c. too little control

44. Some schools hire persons in addition to teachers to help students with special problems. In my opinion, this type of service in this school is

    a. more than enough – it is often forced upon us
    b. enough to help us with our problems
    c. not enough to help us with our problems

45. When a new-comer enters this school, chances are that other students will

    a. welcome him
    b. ignore him
    c. dislike him

46, Homework assignments in this school usually

    a. help us to understand
    b. have little to do with what we learn in class
    c. are just "busy work"

47. In general, teachers in this school pay

    a. too much attention to individual kids and not enough to the class as a whole
    b. not enough attention to individual kids and too much to the class as a whole
    c. about the right attention to individual kids and to the class as a whole

48. In general, my feelings toward school are

    a. very favorable – I like it as it is
    b. somewhat unfavorable – I would like a few changes
    c. somewhat unfavorable – I would like many changes
    d. very unfavorable – I frequently feel that school is pretty much a waste of time

49. In this school the teachers' interest in the students' school work is

    a. not great enough
    b. too great
    c. just about right

APPENDIX B

Counseling Attitude Scale

COUNSELING ATTITUDE SCALE

We are interested in how you feel about the PIANO counseling services available at your school. Your responses to the statements below will indicate your feelings. Read each statement carefully. Decide how you feel about it. Circle the response that best matches your feeling. Please respond to every statement whether or not you have had direct experience with the counseling office.

Example

I enjoy going to the counseling office.

SA A ? D SD

If you *strongly agree*, circle SA

If you *agree*, circle A

If you are *undecided* or *uncertain*, circle ?

If you *disagree*, circle D

If you *strongly disagree*, circle SD

If you have any questions at this time, please raise your hand.

1. I think the counseling facilities are an important part of the school.

   SA A ? D SD

2. I feel the counseling office is unable to solve any kind of problem.

   SA A ? D SD

3. Guidance given to troubled students by counselors is helpful to them.

   SA A ? D SD

4. The efforts made by the counseling staff to help students are not practical.

   SA A ? D SD

5. I believe the counseling office is useful in helping students with their problems.

   SA A ? D SD

6. Talking with counselors at school reduces tension, if nothing else

   SA A ? D SD

7. It is a complete waste of time to go to the counseling office.

   SA A ? D SD

8. I feel the counseling available at school is helpful to students who need it.

   SA A ? D SD

9. It seems to me the counseling office is an important part of the school.

   SA A ? D SD

10. I feel that I can *not* trust anyone at the counseling office to help me.

    SA A ? D SD

11. I regard the counseling office as a place that is serving students with problems of adjustment.

SA A ? D SD

12. I believe the counseling office does *not* adequately interpret test results.

SA A ? D SD

13. The counseling office is a poor excuse for a place where students can take their problems.

SA A ? D SD

14. I believe the counseling office is a good means of helping students with their problems.

SA A ? D SD

15. The counseling office is of no direct help to students. They only tell you what you already know.

SA A ? D SD

16. I think more students should take advantage of the counseling service available at school.

SA A ? D SD

17. I believe the tests used by the counseling office are valuable.

SA A ? D SD

18. I believe the counseling office is *not* interested in students or their problems.

SA A ? D SD

19. I recommend the counseling service at school to all who need help.

SA A ? D SD

20. I feel the school counseling office does enough good work to make it worthwhile.

SA A ? D SD

21. There is a complete lack of organization at the counseling office. You always get the run around.

SA A ? D SD

APPENDIX C

Repeated Measures Analysis of Variance on Attendance Data

The three analyses computed on attendance, disciplinary infractions, and grades were very similar. A general model will be developed here for the analysis of attendance data. In the subsequent appendices the same model will be applied to different sets of data and if variations, are present they will be explained.

All the analyses are four-way repeated measures designs with the repeated measures occuring over one factor; the sampling periods or marking periods. The four factors are sampling periods (A), schools (B), sex (C), and students (D). Students or subjects are usually treated as one factor in repeated measures designs. Students will be considered to be nested under two other factors; schools and sex. The analysis is shown in tabular form in Table 1.

The model for this design is:

$$Y_{ijkl} = \mu + \alpha_l + \beta_k + \gamma_j + \lambda_{i(jk)} + \alpha\beta_{kl}$$

$$+ \alpha\gamma_{jl} + \beta\gamma_{kj} + \alpha\lambda_{il(jk)} + \alpha\beta\gamma_{jkl}$$

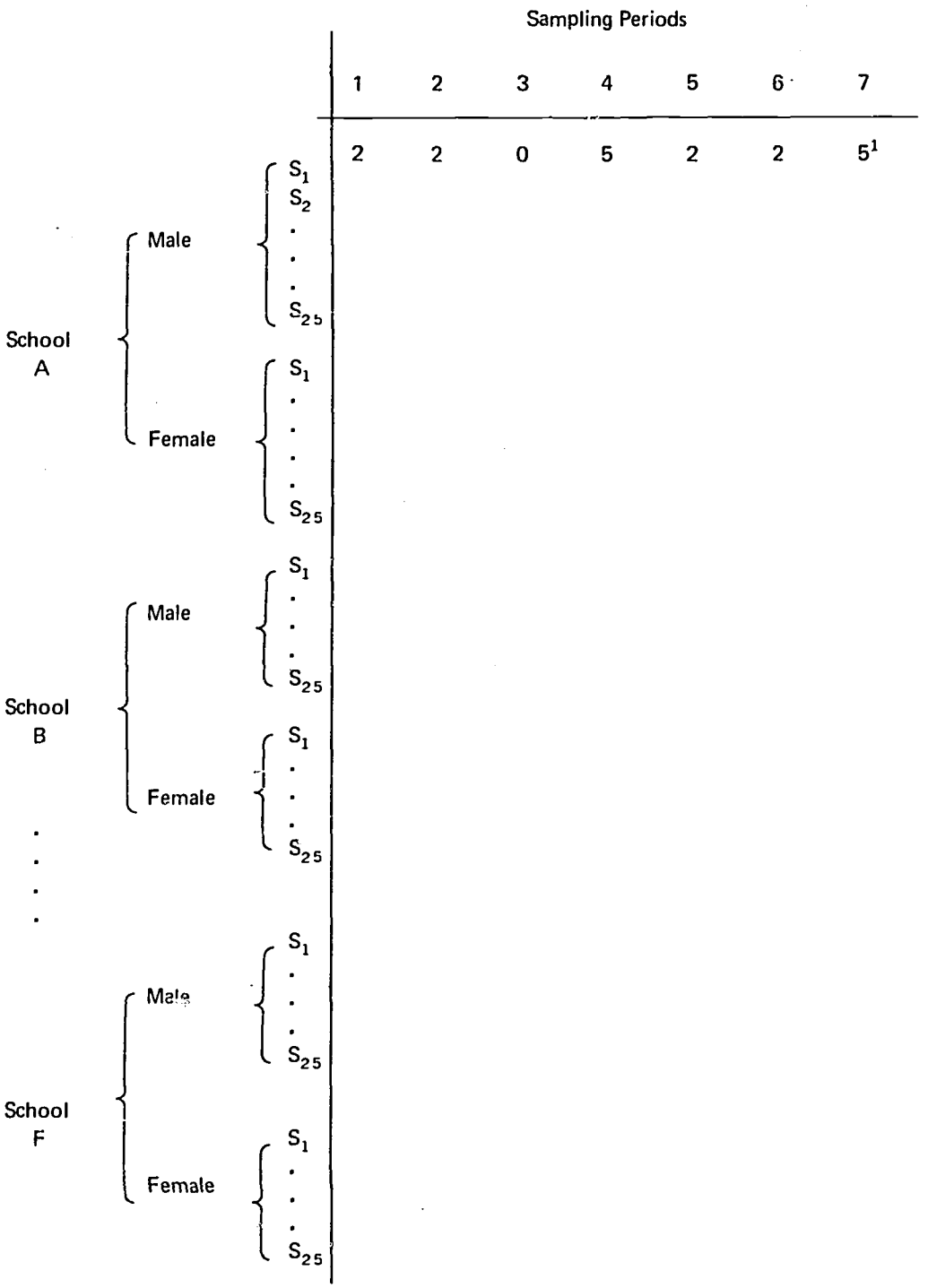
Table 2 presents the variance components for the analysis of variance on the attendance data. It should be noted that there is no provision for an error term in the model because there is only one observation for each cell in the designs. Although the error term is the usual denominator when computing F ratios for an analysis of variance, this is not usually the case for repeated measures on mixed model designs, as will be seen. Several interactions are not included in the model. Interaction effects cannot be computed when one factor is nested under others. Therefore there is no $\beta\lambda$, $\alpha\lambda$, or no three-way interactions including these terms ($\alpha\beta\lambda$, $\alpha\gamma\lambda$) or no four-way interaction.

Cornfield and Tukey (1956) published a general algorithm for determining the expected mean squares for different effects in experimental designs. The expected mean squares are used to decide what denominators are appropriate for each F ratio to be computed. The matrix generated by applying this algorithm is presented in Table 3. Note that the effects included are only those specified in the model.

The variables D_p, D_q, D_r, and D_n, are given values of either 0 or 1 depending on whether the respective factor is considered to be a random or fixed factor. Fixed factors are assigned a value of zero, random factors are one. For the attendance analysis, schools and sex are fixed ($D_q = D_r = 0$); sampling periods and students are considered to be random ($D_p = D_n = 1$). These values are entered into the table to replace the respective variables. D_q and D_r should be set to one in those rows of the table where this effect appears as a nested factor: namely in the rows including the factor represented by λ.

TABLE 1

Schematic Outline of Experimental Design (Attendance Data)



[1]Sample data for attendance was the number of absences during each of the seven five-day sampling periods.

TABLE 2

Analysis of Variance Table of Attendance Data

Effects	Sum of Squares	Degrees of Freedom	Mean Square
Among Students	**2275.26108**	**299**	
Schools (B)	259.58276	5	51.91655
Sex (C)	25.63406	1	25.63046
B x C	56.57523	5	11.31504
Students (S)	1933.47263	288	6.71345
Within Students	**2062.19743**	**1800**	
Sampling Periods (A)	62.18475	6	10.36413
A x B	36.64380	30	1.22146
A x C	6.20952	6	1.03492
A x B x C	34.02475	30	1.13416
A x S	1923.13461	1728	1.11293
TOTAL	**4337.45703**	**2099**	

TABLE 3

Matrix for Determining Expected Mean Squares

Effect	l	k	j	i
α_l	D_p	q	r	n
β_k	p	D_q	r	n
γ_j	p	q	D_r	n
$\lambda_{i\,(jk)}$	p	D_q	D_r	D_n
$\alpha\beta_{kl}$	D_p	D_q	r	n
$\alpha\gamma_{jl}$	D_p	q	D_r	n
$\beta\gamma_{kj}$	p	D_q	D_r	n
$\alpha\lambda_{il\,(jk)}$	D_p	D_q	D_r	D_n
$\alpha\beta\gamma_{jkl}$	D_p	D_q	D_r	n

Table 3 corrected for use with the attendance data appears as Table 4.

TABLE 4

Matrix for Determining Expected Mean Squares
for Attendance Data

Effect	l	k	j	i
1. α_l	1	q	r	n
2. β_k	p	0	r	n
3. γ_j	p	q	0	n
4. $\lambda_{i\,(jk)}$	p	1	1	1
5. $\alpha\beta_{kl}$	1	0	r	n
6. $\alpha\gamma_{jl}$	1	q	0	n
7. $\beta\gamma_{kj}$	p	0	0	n
8. $\alpha\lambda_{il\,(kj)}$	1	1	1	1
9. $\alpha\beta\gamma_{jkl}$	1	0	0	n

The expected mean square for an effect is a weighted sum of every row in the table that includes the subscript (or subscripts) of the effect being evaluated. The weight of each component is the product of all entries in the respective row excluding the entries in the column (or columns) of the subscript (or subscripts) of the effect being evaluated. For example, the mean square for the effect of A must include components from rows 9, 8, 6, 5, and 1 in Table 4 since these effects include the subscript "1" which is associated with α. The term $\alpha\beta\gamma$ is weighted by the product of every entry in row 9 except that in column "1": it is 0. Row 8 becomes simply $\alpha\lambda$ with a weight of one; row 6, 0; row 5, 0; row 1, α with a weight of qrn.

Adding these components together produces:

$$E(MS_A) = \delta^2_{\alpha\lambda} + qrn\ \delta^2_{\alpha}$$

Continuing this process for the remaining effects produce:

$$E(MS_B) = \delta^2_{\alpha} + rn\ \delta^2_{\alpha\beta} + p\ \delta^2_{\lambda} + prn\ \delta^2_{\beta}$$

$$E(MS_C) = \delta^2_{\alpha\lambda} + qn\ \delta^2_{\alpha\gamma} + p\ \delta^2_{\lambda} + pqn\ \delta^2_{\gamma}$$

$$E(MS_S) = \delta^2_{\alpha\lambda} + p\ \delta^2_{\lambda}$$

$$E(MS_{AB}) = \delta^2_{\alpha\lambda} + rn\ \delta^2_{\alpha\beta}$$

$$E(MS_{AC}) = \delta^2_{\alpha\lambda} + qn\ \delta^2_{\alpha\gamma}$$

$$E(MS_{BC}) = n\ \delta^2_{\alpha\beta\gamma} + \delta^2_{\alpha\lambda} + pn\ \delta^2_{\beta\gamma}$$

$$E(MS_{AS}) = \delta^2_{\alpha\lambda}$$

$$E(MS_{ABC}) = n\ \delta^2_{\alpha\beta\gamma}$$

To test each effect an F ratio of expected mean squares is formed such that the denominator contains all terms in the numerator, except the term representing the effect being evaluated. For example, the numerator of the F ratio for the main effect of A or sampling periods is $\delta^2_{\alpha\lambda} + qrn\ \delta^2_{\alpha}$. The last component represents the effect due to sampling periods. The denomintor should be simply $\delta^2_{\alpha\lambda}$. The correct term for the denominator is the interaction term between A and S.

Main Effect A: $$F = \frac{E(MS_A)}{E(MS_{AS})} = \frac{\delta^2_{\alpha\lambda} + prn\ \delta^2_{\alpha}}{\delta^2_{\alpha\lambda}}$$

Main Effect B: (no appropriate denominator found)

Main Effect C: (no appropriate denominator found)

Main Effect S: $$F = \frac{E(MS_S)}{E(MS_{AS})} = \frac{\delta^2_{\alpha\lambda} + p\ \delta^2_{\lambda}}{\delta^2_{\alpha\lambda}}$$

Interactions AB: $$F = \frac{E(MS_{AB}) = \delta^2_{\alpha\lambda} + rn\,\delta^2_{\alpha\beta}}{E(MS_{AS}) \qquad \delta^2_{\alpha\lambda}}$$

AC: $$F = \frac{E(MS_{AC}) = \delta^2_{\alpha\lambda} + qn\,\delta^2_{\alpha\gamma}}{E(MS_{AS}) \qquad \delta^2_{\alpha\lambda}}$$

BC: (no appropriate denominator found)

The interaction of AS and AFC cannot be evaluated because of the lack of an error term in the design.

Although simple denominators are not available to evaluate the effect of B, C, and BC, quasi-F ratios can be used to make these tests (Winer, 1968). Several expected mean squares are combined to produce the necessary terms for a denominator. For example, the denominator for "B" requires the components:

$$\delta^2_{\alpha\lambda} + rn\,\delta^2_{\alpha\beta} + p\,\delta^2_{\lambda}$$

This can be obtained by the combination

$$E(MS_{AB}) + E(MS_S) - E(MS_{AS})$$

Therefore

Main Effect B: $$F = \frac{E(MS_B)}{E(MS_{AB}) + E(MS_S) - E(MS_{AS})}$$

$$= \frac{\delta^2_{\alpha\lambda} + rn\,\delta^2_{\alpha\beta} + p\,\delta^2_{\lambda} + pqn\,\delta^2_{\beta}}{\delta^2_{\alpha\lambda} + rn\,\delta^2_{\alpha\beta} + p\,\delta^2_{\lambda}}$$

Also

Main Effect C: $$F = \frac{E(MS_C)}{E(MS_{AC}) + E(MS_S) - E(MS_{AS})}$$

Interaction BC: $$F = \frac{E(MS_{BC})}{E(MS_{ABC}) + E(MS_{AS})}$$

The results of the computations for each of the evaluated effects are presented in Table 5.

TABLE 5

F Ratios for Testing
Significance of Effects on Attendance Data

Effect	F Ratio	Degrees of Freedom	Value of F Ratio	Approximate Tabled F Value (α=.05)
Sampling Period (A)	MS (A) / MS (AS)	6/1728	9.31*	2.10
Schools (B)	MS (B) / MS (AB) + MS (S) – MS (AS)	5/228	7.61*	2.27
Sex (C)	MS (C) / MS (AC) + MS (S) – MS (AS)	1/197	3.86	3.92
Students (S)	MS (S) / MS (AS)	228/1728	6.03*	1.35
A x B	MS (AB) / MS (AS)	30/1728	1.10	1.46
A x C	MS (AC) / MS (AS)	6/1728	.93	2.10
B x C	MS (BC) / MS (ABC) + MS (AS)	5/980	5.04*	2.23

*Significant at better than the specified .05 level

APPENDIX D

Repeated Measures Analysis of Variance on Disciplinary Infractions

This analysis is very similar to the one for attendance data. Factor A of the analysis becomes sampling periods for disciplinary data which was four periods instead of seven. Otherwise the analysis is the same. Table 1 presents the component of variance and Table 2 the F ratios needed to test the significance of each effect.

TABLE 1

Analysis of Variance Table for
Disciplinary Infractions Data

Effects	Sum of Squares	Degrees of Freedom	Mean Square
Among Students	**2004.68205**	**299**	
Schools (B)	214.58665	5	42.91733
Sex (C)	28.82999	1	28.8299
B x C	1729.91542	5	6.27000
Students (S)	1729.91542	288	6.00665
Within Students	**1097.44907**	**900**	
Sampling Period (A)	14.12667	3	4.70889
A x B	36.47333	15	2.43155
A x C	1.01667	3	0.33889
A x B x C	35.96333	15	2.39756
A x S	1009.86407	864	1.16882
TOTAL	**3102.12612**	**1199**	

TABLE 2

F Ratios for Testing
Significance of Effects on Disciplinary Data

Effect	F Ratio	Degrees of Freedom	Value of F Ratio	Approximate Tabled F Value ($\alpha = .05$)
Sampling Periods (A)	MS (A) / MS (AS)	3/864	4.03*	2.65
Schools (B)	MS (B) / MS (AB) + MS (S) – MS (AS)	5/102	5.90*	2.3
Sex (C)	MS (C) / MS (AC) + MS (S) – MS (AS)	1/164	5.57*	3.9
Students (S)	MS (S) / MS (AS)	288/864	5.14*	1.2
A x B	MS (AB) / MS (AS)	15/864	2.08*	1.68
A x C	MS (AC) / MS (AS)	3/864	.29	2.62
B x C	MS (BC) / MS (ABC) + MS (AS)	5/34	1.76	2.53

*Significant at better than the specified .05 level.

APPENDIX E

Repeated Measures Analysis of Variance on Grades

Although this analysis is quite similar to those presented in the preceeding two appendices, there is one structural change. The sampling periods were not randomly selected by the experimenter but were fixed. The variable (D_p) in the table generaged by the Cornfield-Tukey Algorithm is therefore not I as in the previous designs but is zero (0). The relevant values are presented in Table 1.

TABLE 1

Matrix for Determining Expected
Mean Squares for Grades Data

Effect	l	k	j	i
1. α_l	0	q	r	n
2. β_k	p	0	r	n
3. γ_j	p	q	0	n
4. $\lambda_{i\,(jk)}$	p	1	1	1
5. $\alpha\beta_{kl}$	0	0	r	n
6. $\alpha\gamma_{jl}$	0	q	0	n
7. $\beta\gamma_{kj}$	p	0	0	n
8. $\alpha\lambda_{il\,(kj)}$	0	1	1	1
9. $\alpha\beta\gamma_{jkl}$	0	0	0	n

Only the expected mean squares that are different than those given in Appendix C are listed below.

$E(MS_B) = p\delta^2_{\lambda} + prn\ \delta^2_{\beta} =$

$E(MS_C) = p\delta^2_{\lambda} + pqn\ \delta^2_{\gamma}$

$E(MS_S) = p\delta^2_{\lambda}$

$E(MS_{Bc}) = pn\delta^2_{\beta\gamma}$

The appropriate F ratios for these effects now become:

Main Effect B: F = MS(B) / MS(S)

Main Effect C: F = MS(C) / MS(S)

The components of variance information is included here in Table 2 and the F ratios needed for testing the significance of effects on the Grade data is in Table 3.

TABLE 2

Analysis of Variance Table for Grades Data

Effects	Sum of Squares	Degrees of Freedom	Mean Square
Among Students	**69.64022**	**299**	
Schools (B)	14.13480	5	2.82696
Sex (C)	1.23035	1	1.23035
B x C	0.68282	5	0.13656
Students (S)	53.59225	288	0.18608
Within Students	**15.28232**	**300**	
Grades (A)	1.42984	1	1.42984
A x B	0.37385	5	0.07477
A x C	0.27864	1	0.27864
A x B x C	0.21478	5	0.04396
A x S	12.98521	288	0.04509
TOTAL	**84.92254**	**599**	

TABLE 3

F Ratios for Testing
Significance of Effects on Grades Data

Effect	F Ratio	Degrees of Freedom	Value of F Ratio	Approximate Tabled F Value ($\alpha = .05$)
Marking Periods (A)	MS (A) / MS (AS)	1/288	31.71*	3.86
Schools (B)	MS (B) / MS (S)	5/288	15.19*	2.26
Sex (C)	MS (C) / MS (S)	1/288	6.61*	3.86
Students (S)	------			
A x B	MS (AB) / MS (AS)	5/288	1.66	2.26
A x C	MS (AC) / MS (AS)	1/288	6.18*	3.86
B x C	------			

*Significant at better than the specified .05 level.

APPENDIX F

Standardized Test Results

Use Table 45B for Historical Regression Design

45B. Standardized Test Results

In the table below, please enter the requested information about the tests used to evaluate the effectiveness of major project components/activities in achieving desired objectives. Attach additional sheets if necessary. Before completing this question, read all footnotes.

Component Code	Activity Code	Objective Code	Test Used (MAT, CAT, etc.)	Form	Level	Total N 1/	2/ Group ID	Sample			Pretest		5/ Predicted Posttest Mean 4/	Actual Posttest		7/ df	Statistical Data 6/		
								Size	3/ Y	N	Date	4/ Mean		Date	4/ Mean		Stat Used	Test Obtained Value	Specify Level of Significance Obtained (e.g. p ≤ .05; ≤ .01)
READING			MAT	G		1250	PIANO 9	491		X	4-72 11-72	6.784	7.274	4/73	7.26	490	t	-.213	one tail: p>.4 two tail: p>.8
NOTE: ABOVE REFLECTS SCORES OF FIVE (5) SCHOOLS.																			
READING			MAT GATES-MACGINITIE	G.		1500	PIANO 9	570			4-72 11-72	6.529	6.991	4/73	7.056	569	t	1.272	one tail: p<.25 two tail: p<.5
NOTE: ABOVE REFLECTS SCORES OF ALL SIX (6) SCHOOLS, ONE HAVING																			
	ADMINISTERED			THE		GATES-MACGINITIE					READING		TEST						

1/Total N (total number). Indicate the total number of participants in the component.
2/Group I.D. (group identification). Indicate group, e.g. grade 5; grade 3 control; grade 3 treatment (a control group consists of students selected at the same time that treatment participants were selected and who essentially have the same characteristics as the treatment group. The control group does not take part in the compensatory activity, whereas the treatment group does.)
3/Y/N (yes/no) Is sample representative of universe? The control group does not take part in the compensatory activity, whereas the treatment group does.)
4/Mean. Use grade equivalents unless unavailable from publisher's norms. Specify type of mean used.
5/Predicted posttest. Use only for correlated sampled using "historical" regression procedure.
6/Statistical data. Use test of significance for actual posttest v. predicted posttest where correlated samples are used.
7/d.f. (degrees of freedom). Indicate degrees of freedom used in analysis.
8/Test used and value (e.g., t = 3.85, F = 4.17, etc.). Scores for the same individuals should be included in pre and posttest calculations.

Use Table 45B for Historical Regression Design

45B. Standardized Test Results

In the table below, please enter the requested information about the tests used to evaluate the effectiveness of major project components/activities in achieving desired objectives. Attach additional sheets if necessary. Before completing this question, read all footnotes.

Component Code	Activity Code	Objective Code	Test Used (MAT, CAT, etc.)	Form	Level	Total N 1/	2/ Group ID	Sample			Pretest		5/ Predicted Posttest Mean 4/	Actual Posttest		7/ df	Statistical Data 6/		
								Size	3/ Y	N	Date	4/ Mean		Date	4/ Mean		Stat Used	Test Obtained Value	Specify Level of Significance Obtained (e.g. p≤.05; ≤.01)
MATH			MAT			500	Piano 9	103		X		5.497	5.497	9-72 11-72	6.097	102	t	1.886	one tail p<.05 two tail p<.10
NOTE:		ABOVE REFLECTS SCORES OF ONLY TWO SCHOOLS THAT																	
		PROVIDED DATA REQUESTED.																	

1/Total N (total number). Indicate the total number of participants in the component.

2/Group I.D. (group identification). Indicate group, e.g. grade 5; grade 3 control; grade 3 treatment (a control group consists of students selected at the same time that treatment participants were selected and who essentially have the same characteristics as the treatment group. The control group does not take part in the compensatory activity, whereas the treatment group does.)

3/Y/N (yes/no) Is sample representative of universe? The control group does not take part in the compensatory activity, whereas the treatment group does.)

4/Mean. Use grade equivalents unless unavailable from publisher's norms. Specify type of mean used.

5/Predicted posttest. Use only for correlated sampled using "historical" regression procedure.

6/Statistical data. Use test of significance for actual posttest v. predicted posttest where correlated samples are used.

7/d.f. (degrees of freedom). Indicate degrees of freedom used in analysis.

8/Test used and value (e.g., t = 3.85, F = 4.17, etc.). Scores for the same individuals should be included in pre and posttest calculations.

REFERENCES

Form, A. L. The construction of a scale on attitudes toward counseling, *Journal Counseling Psychology*, 1955, 2, 96-102

Cornfield, J. and Tukey, J.W. Average values of mean squares in factorials, *Annals Mathematical Statistics*, 1956, 27, 907-949

Getzels, J.W. and Jackson, P.W. *Creativity and Intelligence.* New York: Wiley, 1962.

Lahaderne, H. *Adaptation to School Settings – A Study of Children's Attitudes and Classroom Behavior.* ERIC DOCUMENT ED 012 943 March, 1967.

Shaw, M.E. and Wright, J.M. *Scales for the Measurement of Attitudes.* New York: McGraw Hill, 1967.

Winer, B.J. *Statistical Principles in Experimental Designs.* New York: McGraw Hill, 1962.