DOCUMENT RESUME

ED 077 700 SE 016 149

AUTHOR Taylor, Derek B.; Fleming, Margaret
TITLE Individually Prescribed Instruction Program (Mathematics), Disadvantaged Pupil Program Funds, Fund Number 97-19, 1971-72 Evaluation.
INSTITUTION Cleveland Public Schools, Ohio. Div. of Research and Development.
PUB DATE Dec 72
NOTE 69p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Curriculum; Disadvantaged Youth; Elementary School Mathematics; Evaluation; Individualized Instruction; *Instruction; Mathematics Education; *Research; Urban Education
IDENTIFIERS *Individually Prescribed Instruction; Research Reports

ABSTRACT

The third year of a field test of the Individually Prescribed Instruction (IPI) program in two inner-city elementary schools is reported. The results of the Stanford Achievement Tests administered to IPI and Control children in grades two, four, and six showed that IPI students scored higher on one out of three subtests at grade four, on two out of three at grade six, and none out of two at grade two. Results of longitudinal analyses provide no evidence of IPI superiority in mathematics achievement over three years of operation. Cross-sectional analyses provided some evidence of a gradual rise in achievement levels at certain grade levels. Also briefly discussed in the document are financing, staff training program, and school organization. (DT)

INDIVIDUALLY PRESCRIBED INSTRUCTION PROGRAM
(Mathematics)

DISADVANTAGED PUPIL PROGRAM FUNDS

FUND NUMBER 97-19

1971-1972 EVALUATION

(YEAR 3)

Prepared
By

Derek B. Taylor
Research Associate

Margaret Fleming
Directing Supervisor

Cleveland Public Schools

DIVISION OF RESEARCH AND DEVELOPMENT

December, 1972

INDIVIDUALLY PRESCRIBED INSTRUCTION PROGRAM
(Mathematics)

I. INTRODUCTION

A. Needs and Rationale

From among the new instructional methodologies emerging from research in education, the Cleveland Public Schools selected the Individually Prescribed Instruction (IPI) approach for experimental trials in teaching mathematics in two Title I elementary schools. Implementation of the program began in the fall of 1969 and has continued through the 1971-72 school year. This report is the evaluation of the third year of program operation.

The IPI technique is being installed in school systems across the nation by Research for Better Schools (RBS), a federally funded regional education laboratory. IPI is designed to provide a program of study in math that is individually tailored for each student on each day. The program has six elements that distinguish it from conventional school procedures. They are:

- Detailed specifications of educational objectives.
- Organization of methods and materials to attain these objectives.
- Careful determination of each pupil's present competence in a given subject.
- Individual daily evaluation and guidance of each pupil.
- Provision of frequent monitoring of student performance in order to inform both the pupil and the teacher of progress toward an objective.
- Continual evaluation and strengthening of the curriculum and instructional procedures.

1. Objectives

Generally, this program seeks to implement a new instructional process in mathematics for Grades 1-6. The

objectives fall into two broad categories: (1) process objectives related to the implementation of the program in the target schools, and (2) product objectives related to the expected results of the successful implementation of the program.

a. Process Objectives

(1) To instruct target school staff during the summer of 1972 in individualizing reading and spelling instruction by means of a training program in Individually Prescribed Instruction designed and conducted by Research for Better Schools.

(2) To provide an individualized program of mathematics study for each pupil enrolled at the target schools by implementation of the Individually Prescribed Instruction Program according to the format specified by Research for Better Schools.

b. Product Objective

(1) To increase the level of achievement of pupils in mathematics significantly beyond the level demonstrated by pupils in two Control schools as measured by the pre-post administration of standardized mathematics achievement tests.

B. Historical Background

The IPI system was developed by the Learning Research and Development Center of the University of Pittsburg. and is being field tested in school systems across the country by Research for Better Schools. The introduction of the IPI program into the Cleveland Public Schools grew out of visits by representatives of the Department of Mathematics and the Division of Research and Development to other school systems where IPI programs were in operation. Training of the principals of the target schools took place in April, 1969 and of the teaching staff in June and July, 1969. Implementation of the program in the Cleveland Public Schools began in September, 1969.

The evaluation of Years 1 and 2 of program operation resulted in the following findings:

1. All staff training was conducted as proposed and according to the format established by RBS.

2. A team of neutral observers rated the program operation as well-organized and implemented. Suggestions were made for a greater quantity and variety of manipulative materials for use in instruction and for greater variety in downtime assignments.

3. An independent evaluation conducted by RBS in Year 1 indicated that both schools showed a high degree of accuracy in correctly following the IPI procedures.

4. Comparison of the performance on standardized achievement tests showed the following:

   In Year 1 the children in IPI schools performed significantly better than children in Control schools on one out of three subtests at Grade 5. At Grades 1 and 3 the children in Control schools performed better than children in IPI schools to a degree that reached or approached statistical significance on all arithmetic subtests

   In Year 2 the children at IPI schools performed significantly better than children in Control schools on the one standardized arithmetic test administered at Grade 1. At Grade 3 children in Control schools scored significantly higher on one of the two tests administered. At Grade 5 Control children scored significantly higher on one out of three tests administered.

5. Longitudinal analyses comparing children who had been in the IPI or Control schools for the two full years of operation showed only one significant difference in math performance on standardized tests. Fifth graders who had been in Control schools for two years scored significantly higher than IPI children on the Computations subtest.

6. Cross sectional analyses of changes in performance at given grade levels across time indicated that the IPI Program had little effect on performance levels over two years of operation.

C. Summary of Program Operations in Year 3

1. Participants

In Year 3 the IPI Program operated in the same two elementary schools as in Years 1 and 2, Parkwood and Stephen E. Howe. All students in Grades 1-6 received the services of the program. Parkwood served 437 pupils and Howe served 446 for a total of 883 children receiving project services.

During the first year of program operation (1969-70) approximately $137,210 was spent on the IPI Program in addition to General Fund expenditures in the two schools served. By Year 3 IPI expenditures had increased to approximately $185,641. Per pupil expenditure data are presented in Table 1.

Table 1

Per Pupil Expenditures from IPI and General Funds in Year 1, Year 2 and Year 3

	Per Pupil Expenditure		
Program Year	IPI	General Fund*	Total
Year 1	$121	$52*	$173
Year 2	$162	$56	$218
Year 3	$210	$64	$274

*General Fund expenditures for math instruction are based on an estimated 12% of the total General Fund per pupil expenditure.

The table shows that over the course of three years the per pupil expenditure from IPI funds increased 74%, general fund expenditures increased by 23% and the total per pupil expenditures in the project schools increased by 58%. In Year 3 the addition of IPI funds to the general funds normally spent

on math instruction increased the per pupil expenditure in the project schools by 22%.

D. Questions to be Answered by Evaluation

The evaluation of Year 3 of the IPI Program was addressed to four questions based on the objectives of the program:

1. Did the project staff complete the training program for IPI (Process Objective 1)
2. How were the IPI schools organized for implementation of the program? (Process Objective 2)
3. Has the product objective related to changes in achievement level attained? (Product Objective 1)
4. What has been the impact on achievement levels in math over the three years of operation of the program?

II. HIGHLIGHTS OF FINDINGS

A. Summary of Findings

1. Evaluation Question 1: Did the Project Staff Complete the Training Program for IPI?

FINDING: Staff training in IPI Reading and Spelling did not take place because of a decision not to implement an IPI program in these areas. Staff new to the IPI schools received training IPI Math. The process objective related to staff training was considered achieved.

2. Evaluation Question 2: How Were the IPI Schools Organized for Implementation of the Program?

FINDING: The organization and procedures in the IPI schools were found to be essentially the same as in previous years when a detailed analysis showed that the IPI approach was being implemented according to specification. The process objective related to implementation of program procedures was considered achieved.

a. All staff and materials for the program were provided as planned, although Stephen Howe school reported difficulty in getting orders for IPI materials correctly filled. Materials for the program were maintained in a centrally located materials center as proposed.

b. In both schools the math classes were scheduled and organized to permit the presence of the homeroom teacher, a consultant teacher and two aides. Some classes had the services of two consultant teachers.

c. In both schools weekly planning meetings were held among teachers from the same or adjacent grade levels.

d. In both schools prescription writing was done during class time, bell time, or during free periods. Prescriptions were written by both the homeroom teachers and the consultant teachers. The assignment and scheduling of staff was such that prescriptions were always ready by the beginning of the math period.

e. The staff in both schools used little of the budgeted after-school in-service time, but used planning meetings and regular faculty meetings to accomplish the same goals. In Stephen Howe part of the in-service time was used to construct manuals of remedial or enrichment assignments for children to work on during the time their IPI worksheets were being checked.

f. Both schools made efforts to educate parents in the operation of the IPI Program through PTA meetings, demonstrations, brochures and the use of a unique report card.

3. Evaluation Question 3: Was the Product Objective Related to Changes in Achievement Levels Attained?

FINDING: The achievement objective of IPI superiority over Control schools in performance on standardized math achievement tests was attained at one of the three grade levels tested.

a. Product Objective: IPI Superiority in Achievement in Math

A summary of the results of the Stanford Achievement Tests administered to IPI and Control children in Grades 2, 4 and 6 appears in Table A.

Table A

Summary of Results of Achievement Testing in Year 3

Grade	SUPERIOR GROUP ON EACH MATH SUBTEST		
	Arithmetic Computations	Arithmetic Concepts	Arithmetic Applications
2	None	None	---
4	Control	Control	IPI
6	Control	IPI	IPI

IPI children scored higher than Control children on one out of three subtests at Grade 4, two out of three subtests at Grade 6 and none out of two subtests at Grade 2. The product objective is considered achieved at Grade 6. The performance in Year 3 is somewhat better than in the previous year when IPI children scored higher on only one out of one subtests administered at one of the three grade levels tested.

b. Progress in the IPI Continuum

A pre and post analysis of the percentage of children working at various skill levels in the IPI continuum showed that children at different grade levels progressed through the continuum at approximately the same rate.

4. Evaluation Question 4: What has been the Impact on Achievement Levels in Math over the Three Years of Operation of the Program?

FINDING: The results of longitudinal analyses provide no evidence of IPI superiority in math achievement over

three years of operation. Cross sectional analyses provided some evidence of a gradual rise in achievement levels at certain grade levels.

a. Longitudinal Analysis

Samples of children were selected who had been in the IPI and Control schools for the entire three years of operation of the program. These children were compared for differences in achievement levels at the end of the three years. Three samples of children were tested:

- Grade 1-2 sample - Children entered program in Grade 1, Year 2 and were compared in Grade 2, Year 3.
- Grade 2-3-4 sample - Children entered program in Grade 2, Year 1 and were compared in Grade 4, Year 3.
- Grade 4-5-6 sample - Children entered program in Grade 4, Year 1 and were compared in Grade 6, Year 3.

Table B shows a summary of the results of the achievement comparisons made between IPI and Control children in the longitudinal samples.

Table B

Summary of Results of Longitudinal Achievement Comparisons Between IPI and Control Children

Longitudinal Sample	SUPERIOR GROUP ON EACH MATH SUBTEST		
	Arithmetic Computations	Arithmetic Concepts	Arithmetic Applications
Gr. 1-2	None	None	---
Gr. 2-3-4	Control	None	None
Gr. 4-5-6	Control	Control	None

Of the three longitudinal samples tested, none showed any significant differences in favor of IPI. These results are similar to those in the previous year of operation.

b. Cross Sectional Analysis

Analysis of the changes in performance levels across three years at given grade levels in IPI schools revealed the following results:

(1) Of the three grade levels examined (Grades 3, 5 and 6) only at Grade 6 was there any evidence indicating a possible salutary effect of the IPI Program on achievement levels. At Grade 6, performance levels on the Concepts and Applications subtests appeared to be gradually rising in IPI schools in the face of generally declining performance in other schools. In Grade 3, changes in performance in IPI schools were generally paralleled by similar changes in other schools, and in Grade 5 the performance in IPI schools declined relative to other schools.

B. Implications and Recommendations

1. Discussion of Results

The results of the third year of operation of the IPI Program have shown that (as far as can be determined without a detailed analysis of specific diagnostic and prescription writing procedures) the IPI Program continues to be implemented according to plan. The question that remains is the effectiveness of the program in raising the math achievement levels of the children served. The pre-post analysis of progress through the skill levels within the IPI continuum showed that the children are moving along the continuum, but because of the lack of norms against which to compare their progress, it is difficult to determine whether they are progressing at a "normal" rate. The evaluator must, therefore, rely on information from standardized achievement tests to draw conclusions about the impact of the program on math performance.

The achievement data from the third year of operation showed that the children in IPI schools fared somewhat better in comparisons with children in Control schools than they did in previous years. The results, however, did not indicate a strong impact, except at Grade 6. Similarly, cross sectional

analyses of changes in performance levels at given grade levels across the three years of operation showed little evidence of improvement except at Grade 6. Longitudinal analyses of the performance of only those children who had been in the program continuously for two or three years showed no evidence at all of superior performance by IPI children.

As discussed in the evaluation report for Year 2 of operation, there are several possible reasons why a greater impact on achievement has not been evident. The first possibility is, of course, that the IPI approach simply does not work any better than the traditional methods of math instruction. Other possibilities exist, however. It is possible that the standardized achievement results do not adequately measure what is taught in the IPI continuum, (a possibility that raises questions as to what is taught in the IPI continuum). One way of investigating this thesis is to analyze the standardized tests item by item to determine whether the questions reflect the skills and concepts covered in IPI. The team of consultant teachers in Stephen Howe school was assigned this task, and their analysis of the arithmetic subtests of the Stanford Primary I, Stanford Primary II, Stanford Intermediate I and Stanford Intermediate II achievement tests showed that an average of 94% of the items on these tests were covered in the IPI continuum. The range of correspondence between the Stanford tests and the IPI continuum was from 82% to 100% for nine arithmetic subtests. It would appear that the standardized tests do measure the same things that are taught in the IPI continuum.

It is possible, of course, even though the standardized tests measure basically the same skills as are taught in IPI, that the language of the tests is sufficiently different from the language used in IPI to impair performance on the tests. For example, one o[illegible] the Stanford Intermediate I Arithmetic Concepts makes use of the terms multiplier, multiplicand and minuend, but these terms are not used at all in IPI. To investigate this potential problem, the consultant teachers in the IPI schools constructed achievement tests directly from the IPI materials. Tests were constructed for Grades 3 and 5, and items were selected that, in the judgment of the IPI staff, reflected the instructional activities in which all or most of the pupils had participated. Since the test items reflected what most of the children had been taught in the IPI continuum, a mastery level performance was expected. It was anticipated that each item would be answered correctly by at least 75% of the children. The results showed, however, that no more than 36% of the items on any of these tests were correctly answered by 75% or more of the children. In fact no more than 57% of the items on any of the tests were correctly answered by more than 60% of the children. These results are not conclusive, because these IPI tests are as yet in unrefined form, but they do suggest that lack of evidence of improvement in math achievement may not be due to the use of standardized tests.

Another possibility for explaining the apparent lack of progress in raising achievement levels is that the IPI

system teaches the math skills that the Cleveland Schools wants taught, but that it teaches them in a different sequence from that normally followed. That is, by the time he leaves elementary school, a child in IPI may have covered the same concepts and skills as a child not in IPI (hopefully with greater mastery), but at any given point in time during their elementary school careers, they may be working on entirely different skills. If this is the case, the only valid assessment of IPI possible is a longitudinal study in which achievement levels of IPI and Control children are compared only after the IPI children have spent all or most of their elementary school years working in the IPI system. One problem with this approach, however, is that the percentage of children remaining in the schools that long is relatively small. Less than one third of the children present in the IPI schools during Year 1 were still in the schools by the end of Year 3. Therefore, the whole question of whether the IPI Program might have an impact on achievement if children participated for their entire elementary school tenure may be academic.

Finally, despite the indications that the program guidelines have been faithfully followed by the IPI staff, over the course of three years it is possible that deviations from the precise testing and prescription criteria established by the program's designers have crept into the operation. If so, these deviations may be reflected in lack of improvement in math achievement. It was recommended in the evaluation report

for Year 2 that another evaluation of operations accuracy be conducted by Research for Better Schools, but that organization has reported that they no longer perform that service. Self-monitoring instruments have been developed, however, and their use in the Cleveland IPI schools may yield valuable information as to the accuracy with which the IPI procedures are being followed.

In the long run, the decision about the future of the IPI Program will have to balance program costs against program benefits. There is some evidence that achievement levels in math in the IPI schools may be improving slightly at certain grade levels. But is is also clear that over the course of three years the cost of the program has risen 74% to $210 per pupil. The question becomes whether the present or future gains in achievement are great enough to justify the expense, or whether more efficient and less costly avenues to the same ends can be found.

At any rate, the results of field tests of experimental programs in the schools must be interpreted in light of the number of influences over which the evaluator has no control; experimental treatment populations change, control populations change, samples may become biased through non-random attrition. All of these factors may cause changes in performance that have nothing to do with the experimental treatment under investigation. Evaluation must, therefore, take a long and cautious view of results and data must be examined over a period of several years before administrators can state with some degree of confidence that a given program is or is not effective.

2. Recommendations

a. It is recommended that the IPI Program be continued in the same two schools in which it presently operates for the next one or two years until sufficient data are available for an accurate appraisal of long term program effects.

b. It is recommended that a decision on expansion of the IPI Program to other schools be deferred until the results of a long term evaluation are available.

c. It is recommended that the development and refinement of achievement tests constructed from the IPI materials proceed.

d. It is recommended that self-monitoring procedures be instituted in each IPI school, using the instruments developed for this purpose by RBS.

III. PROJECT DESCRIPTION

A. Participant Characteristics

Students in both Parkwood and Stephen E. Howe schools have exhibited a pattern of increasing deficit in achievement in math as they progress from Grade 1 through Grade 6. Before the IPI Program was begun, the average achievement level in math was over one year below grade level norms by the end of the sixth grade. Both schools are "poverty schools" with an average of 42% of the children on public assistance. Pupil mobility rates are high, averaging 70% during the school year 1971-72. Table 2 shows the number of students served in the project by grade level. (See Appendix A for a breakdown by school).

Table 2

Number of Pupils Served by IPI Program by Grade Level

Grade Level		Number of Children
EMR	-	18
1	-	150
2	-	168
3	-	147
4	-	127
5	-	117
6	-	156
	Total	883

B. Project Operations (Program Plan)

1. Organizational Details and Activities

The basic plan for IPI is as follows:

a. The subject area of mathematics is structured into a hierarchical continuum of several hundred learning objectives along which the student progresses at his own pace. The objectives are sequenced so that before a student begins to study a new mathematical process, he has already mastered those skills that are prerequisite to learning the new process. Thirteen areas of mathematics are covered:

Numeration
Place Value
Addition
Subtraction
Multiplication
Division
Combination of Processes
Fractions
Money
Time
Systems of Measurement
Geometry
Special Topics

Each of these areas is divided into skill levels A through G, each skill level involving progressively more complex processes.

b. Placement tests are given when a child enters the program to determine his strengths and weaknesses and to determine at what level he will begin working in each area of math. At the beginning of each unit a pre-test is given to identify weaknesses in that unit, and then the child begins work in the Standard Teaching Sequence (STS) leading him through the pre-determined sequence of objectives. Curriculum Embedded Tests (CETs) are given as part of the STS and serve to test skill mastery within a unit. At the completion of a unit, a post test is given to confirm mastery of the entire unit.

c. The teacher, on the basis of the test information, writes learning prescriptions for each child, detailing his work assignments. These may most often include assignments in the STS but may also include teacher constructed worksheets, textbook assignments, manipulatives, etc. All learning materials are kept in a centrally located material center staffed by aides.

d. A folder is maintained on each child containing his prescriptions, current work assignment and his record of progress. At the beginning of each class, the child takes his folder, reads his prescription for the day, and proceeds to the materials center to get the numbered worksheets dealing with his objectives for the day. Or, he may continue working on the previous day's prescription. Most assignments consist of the child working alone on his individual assignment, but the teacher may also prescribe individual tutoring, small group work or large group work.

e. As a pupil completes sections of his work assignment he, turns them in to a teacher aide who checks and returns them. The period during which a pupil's worksheets are

being checked is called "downtime". After his work has been checked, the pupil either asks for a new prescription from the teacher or asks the teacher for assistance. Pupils request new prescriptions or assistance from the teacher by displaying a red flag on their desk.

f. The maintenance of the materials center and all checking of papers are the province of aides, freeing the teacher for teaching, evaluation, diagnosis, and prescription writing. Floating teachers without homeroom responsibilities are provided to increase the number of professional staff working directly with the children.

2. Staffing

The program plans called for the addition of several staff to each project school in order to implement the IPI procedures. Included were ten consultant teachers whose responsibilities were to work with the homeroom teachers in preparing prescriptions and in diagnosing and evaluating student progress. Consultant teachers were also to be responsible for supervising the teacher aides. Eighteen teacher aides were proposed to handle clerical work involved in correcting students' worksheets, to maintain the IPI materials center, and to maintain project records and individual student files.

3. In-Service Training

Program plans provided 30 hours of in-service training in the IPI math technique for teachers and aides new to the project schools. In addition, the entire staff of one IPI school was to receive 50 hours of training during the summer of 1972 for implementation of an IPI Reading and Spelling program.

Six after school inservice meetings for IPI staff were planned for review of curriculum changes planned by RBS and for the development of IPI-related instructional materials.

4. Parent Involvement

Because of the emphasis in IPI upon the mastery of behavioral objectives within a math continuum rather than the traditional marking system, communication with the parents about the program was considered important by program planners. Plans for increasing parent understanding of the program included parent meetings, classroom demonstrations and displays, a unique report card system, and packets of literature describing the program.

IV. EVALUATION

Evaluation activities during Year 3 focused primarily on assessment of the product objective related to achievement levels. Although data related to success in achieving the process objectives were collected, the smooth implementation of procedures during Years 1 and 2 suggested that less emphasis on assessment of implementation was required.

A. Basic Design

The evaluation design employed either a status description or a cross-nested experimental-control model, depending on the nature of the objectives under consideration. The status description was built on the results of interviews with project personnel and the examination of school and project records. For experimental-control comparisons, two Control schools (Hough and Hazeldell) were selected to match the IPI schools as closely as

possible on percentage of students on public assistance, attendance, mobility, achievement levels, intelligence levels, and racial composition. These Control schools received the services of other Title I and DPPF projects. Appendix B lists the characteristics on which the IPI and Control schools were originally matched and shows the degree of matching attained.

B. Presentation of Findings

The evaluation procedures were addressed to four basic questions about project operations and results, and the findings will be organized in the same way. Each basic evaluation question will be posed, followed by an examination of the data collected on the objectives that are pertinent to that evaluation question. Included where necessary will be a description of instrumentation and data collection procedures.

EVALUATION QUESTION 1

DID THE PROJECT STAFF COMPLETE THE TRAINING PROGRAM FOR IPI?

1. Data Collection

Data related to Evaluation Question 1 were collected through interviews with the principals of the project schools.

2. Process Objective 1: Staff Training in IPI Reading and Spelling

The original program plans for Year 3 included training the staff of one of the IPI schools in IPI techniques for reading and spelling with the aim of implementing a reading and spelling program in the fall of 1972. However, these plans were postponed after the evaluation of Year 2 indicated

that a more complete investigation of the IPI math program would be required for a decision on the effectiveness of the IPI approach in general. Therefore the training in IPI Reading and Spelling did not take place.

3. Staff Retraining in IPI Math

Staff turnover and additions required the training of several new teachers and aides in the IPI technique. In both schools, five full-day training sessions for new staff were conducted during the week before school. In Stephen Howe school, where both teachers and aides were in need of training, the sessions were conducted by the consultant teachers. In Parkwood, where only one aide was in need of training, the sessions were conducted by the head teacher aide.

The teachers were trained using a teacher-training curriculum devised by RBS. Materials included behavioral objectives, self instructional materials, and recommended learning settings. Training activities included concept building related to IPI, an analysis and application of the concepts to IPI, practice in using IPI skills and materials as routine exercises, and discussion designed to provide an opportunity for clarification and expression of reaction to IPI. The training sessions were designed to equip the teacher to conceptualize a model of IPI as a basis for instructional decision making and to plan and conduct IPI in the classroom.

Process Objective 1 is considered achieved.

EVALUATION QUESTION 2

HOW WERE THE IPI SCHOOLS ORGANIZED FOR IMPLEMENTATION OF THE PROGRAM?

1. Data Collection

Data related to Evaluation Question 2 were collected through interviews with the principals of the project schools.

2. Process Objective 2: To Provide an Individualized Program of Math Instruction through Implementation of IPI Procedures.

a. Materials

Both IPI schools maintained a materials center where testing and prescription materials required by the program were stocked. The materials centers were under the continuous supervision of a teacher aide. At Stephen Howe considerable difficulty was reported in keeping the materials center stocked with the needed IPI materials. The difficulty lay in orders for materials being slowly and incorrectly filled by RBS. The problem was compounded by changes in the curriculum materials instituted by RBS. These changes required wholesale replacement of large volumns of STS booklets.

b. Staff

Five consultant teachers without homeroom assignments were assigned to each school, an increase of one per school over Year 2. These teachers worked exclusively in the area of IPI, rotating among classrooms to help the homeroom teachers in evaluation, diagnosis and prescription writing during the IPI math periods.

Both schools were also assigned teacher aides for work exclusively in the IPI Program. Parkwood received the services of eight aides and Stephen Howe was assigned ten. The aides were responsible for all record keeping, distribution and maintenance of materials and supplies, the operation of the materials center, and scoring the students' worksheets.

c. Organization and Scheduling

In both schools the lack of reading skills hindered the implementation of the program with children in Grade 1. In Stephen Howe an intensive reading vocabulary program geared specifically toward the language used in the IPI materials was held during the first week of school. This vocabulary program made it possible for Grade 1 children to begin work in the IPI continuum in September. In Parkwood, Grade 1 children began working in IPI in January.

In both schools the scheduling was arranged so that each math class had the services of the homeroom teacher, a consultant teacher and two aides. In both schools some classes had the services of two consultant teachers. In Stephen Howe two consultant teachers worked in each classroom in Grades 1 and 2. In Parkwood two consultant teachers worked in the largest classes, irrespective of grade level.

d. Planning

In both schools IPI meetings among homeroom teachers were held regularly once a week to discuss problems, share ideas, plan downtime work, review progress, and select children with similar problems for small group instruction. In both schools the meetings involved teachers from the same or adjacent grade levels.

In both schools the planning meeting day for the teachers was "seminar day" for the children. While the homeroom teachers met, their classes were covered by the consultant teachers who conducted seminars to introduce new IPI concepts to the children.

e. Prescription Writing

At both schools prescriptions for childrens' work were written by homeroom and consultant teachers during class time, free periods and during bell time. There was no special time set aside specifically for writing prescriptions. If prescription writing was not completed by the end of the math period, the standard practice in both schools was for a consultant teacher to finish it later in the day to insure that all prescriptions were ready in time for the math period the next day. This practice was made possible by the addition of one extra consultant teacher in each school in Year 3.

f. In-Service Meetings

Aside from the in-service training for new teachers and aides that was discussed under Evaluation Question 1, money was budgeted for six after school in-service meetings in each school. The purpose of the meetings was to review curriculum changes planned by RBS and to develop IPI-related instructional materials.

Both schools conducted just one such after-school meeting, but used the regularly scheduled planning and faculty meetings to accomplish the same ends. In Stephen Howe part of the in-service time was used to construct manuals of activities

for teachers to assign students during "downtime", the period during which students waited for their worksheets to be checked. Four manuals were developed, one for grades 1 and 2, one each for grades 3 and 4 and one for grades 5 and 6. The manuals were designed to provide either enrichment or remedial assignments related to the area and skill level of the math continuum in which the student was working.

g. Parent Involvement

Both schools continued efforts to communicate the nature and purpose of the IPI Program to the parents of the children served. These efforts included demonstrations of the IPI technique to PTA meetings and groups of parents visiting the classrooms, distribution of brochures describing the program, and the use of a report card that shows where the child placed in the math continuum and explains his progress.

h. Summary of Evaluation Question 2: How Were the IPI Schools Organized for Implementation of the Program?

The results of the first two years of operation of the IPI Program showed that both schools were accurately following the IPI procedures specified by RBS. This conclusion was derived from both local evaluation data and the results of a process evaluation conducted by RBS. Given these results during the first two years, it was not considered necessary to examine the process aspects of the program in such depth during the third year. Since the data that were collected during the third year indicate that essentially the same organization was maintained as during the first two years, it seems safe to assume that implementation of the IPI Program continued to accurately reflect the procedures specified by RBS. Process Objective 2 is considered achieved.

EVALUATION QUESTION 3

WAS THE PRODUCT OBJECTIVE RELATED TO CHANGES IN ACHIEVEMENT LEVELS ATTAINED?

1. Product Objective: IPI Superiority over Control Schools in Math Achievement

a. Data Collection

In order to assess the impact of the IPI Program on achievement levels in basic math skills, comparisons were made between IPI and Control schools on performance on the arithmetic subtests of the Stanford Achievement Tests. The tests were administered to a sample of children from Grades 2, 4 and 6 in IPI and Control schools in September, 1971 and again in May, 1972.

Data were collected from students in regular classes only. Students in any kind of special class such as Major Work, Enrichment, Listening Post or EMR were not tested. Table 3 shows the level and form of the test administered to each grade level and the number of students included in the analysis of the data.

Table 3

Form and Level of Stanford Achievement Test Administered to Grades 2, 4 and 6 in IPI and Control Schools, and Size of Sample Tested

Grade Level	Sample Size		Level and Form of Test
	IPI	Control	
2	108	111	Pre - Primary I, Form X Post - Primary II, Form W
4	97	121	Pre - Intermediate I, Form X Post - Intermediate I, Form W
6	125	210	Pre - Intermediate II, Form X Post - Intermediate II, Form W

Data analysis was performed on the scores of only those children for whom both pre and post test scores were obtained. Consequently, the results reflect the achievement of only those students who attended IPI or Control schools for the nine-month period from September, 1971 to May, 1972.

Data analysis was performed on the raw scores attained on the tests, i.e., the number of correct responses. Raw scores on the Stanford Achievement Tests are often transformed into grade equivalent scores to show a student's performance relative to the national norms. Grade equivalents are very useful for descriptive purposes, but because they are not an equal-interval scale, raw scores were preferred for statistical analysis.

IPI and Control schools were compared with respect to post-test math achievement scores. To avoid the problem of the post score differences being merely a reflection of initial differences in pre scores, the data were analyzed by means of multivariate analysis of covariance (MANCOVA). MANCOVA determines the effects that selected variables (covariates) may have on post scores and adjusts the data for these effects before making comparisons. In the present analysis the students' pre-test math achievement scores and their PLR scores were used as covariates, and the IPI and Control schools were compared on post-test performance after adjusting for the effects of these covariates.

b. Achievement Results - IPI vs. Control

Table 3 shows the results of comparisons between the post-test math performance of IPI and Control children in Grades 2, 4 and 6. In the analysis, math pre-test scores and PLR scores were used as covariates. Pre-test scores were adjusted for unequal N, and post scores were adjusted for unequal N and the effects of the covariates. Full statistical data appear in Appendix C.

Table 3

Mean Adjusted Post Arithmetic Subtest Scores for Children in Grades 2, 4 and 6 in IPI and Control Schools

		MEAN ADJUSTED ARITHMETIC SUBTEST SCORES					
		Computations		Concepts		Applications	
Grade	Group	Post Score	Superior Group	Post Score	Superior Group	Post Score	Superior Group
2	IPI	21.2	None	16.3	None	---	---
	Control	21.8		16.9		---	
4	IPI	12.8	Control*	13.4	Control***	11.2	IPI****
	Control	16.6		14.6		10.4	
6	IPI	13.7	Control*	12.3	IPI**	15.2	IPI*
	Control	16.3		11.4		13.2	

*$p < .0001$
**$p < .02$
***$p < .06$
****$p < .095$

The table shows that during Year 3 the IPI children scored higher than Control children on the Concepts subtest at Grade 6 and on the Applications subtest at Grades 4 and 6. The difference on the Applications subtest was statistically significant at Grade 6 and approached significance at Grade 4. The Control children scored significantly higher than the IPI children on the Computations subtest at Grades 4 and 6.

On the Concepts subtest Control children out scored IPI children at Grade 4 to a degree that approached statistical significance. These results do not show strong and consistent differences in favor of IPI, although they are somewhat improved over the results from the previous year. In the previous year, significant differences in favor of IPI were found on only one subtest at one grade level. The present results must be characterized, however, as scattered and inconclusive, despite the improvement over Year 2. What IPI strength there was appeared primarily at Grade 6.

At Grade 2 a significant Sex x Treatment interaction was obtained on the Computations subtest. The interaction effect is illustrated in Figure 1. The interaction showed that the lack of a significant difference between IPI and Control children was due to the differential performance between boys and girls in the Control schools. (Full statistical data on the interaction effect are available in Appendix C.)

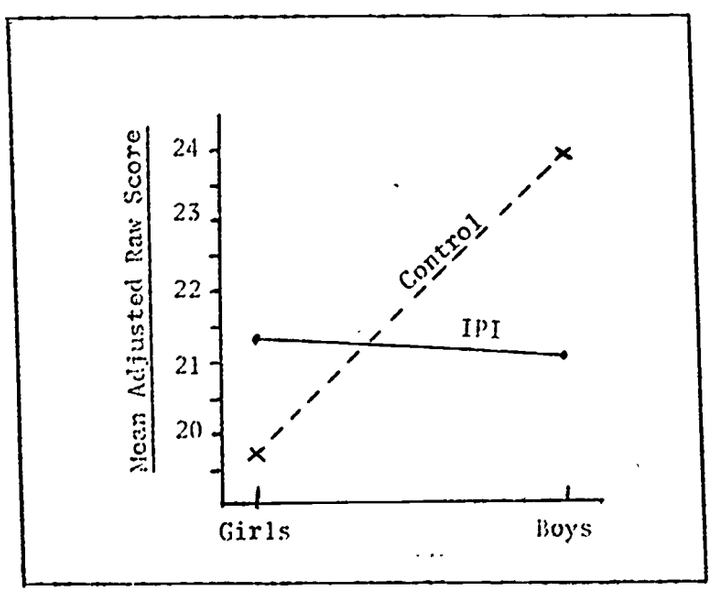

Fig. 1 Sex x Treatment Interaction on Computation Subtest at Grade 2.

In order to show the math performance of the IPI and Control children in relation to grade level norms, the pre and post raw scores were transformed into grade equivalent scores and are presented in graphic form in Figure 2. The results for the three grade levels are plotted on the same scale so that the relationship among performance and the norms will be comparable among grade levels. In reading the figure, two points should be kept in mind: Grade equivalent scores are not an exact reflection of the raw score obtained on the test and at times will not accurately reflect differences in raw scores. Further, the grade equivalent scores used in Figure 2 were computed from the observed raw score means, not from the means used in the statistical analysis, which were adjusted for the effects of unequal N and covariates.

Figure 2 shows that at Grade 2 both IPI and Control children were performing at or close to the norm level on both math tests. In Grades 4 and 6, however, the IPI children were functioning a year and sometimes up to two years below the norm with little evidence of progress toward the norm during the year. The exception was on the Concepts subtest at Grade 4 where IPI children were eight months below the norm at the pre-test and only three months below at the post-test. Overall, however, it is evident that there was a performance deficit that increased as a function of grade level.

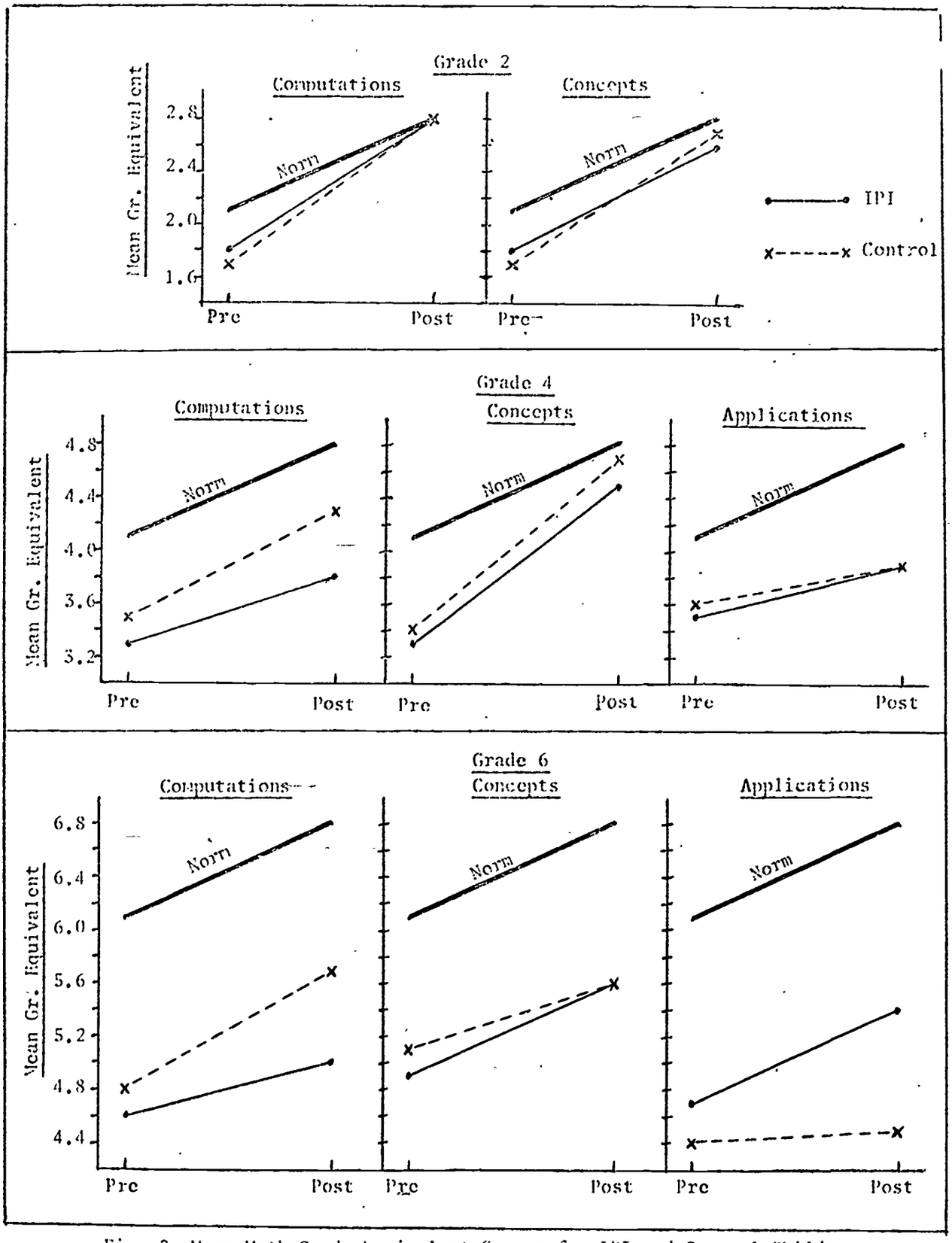

Fig. 2 Mean Math Grade Equivalent Scores for IPI and Control Children in Grades 2, 4 and 6 in Year 3.

c. Achievement Results - Boys vs. Girls

Examination of the math achievement results for sex differences revealed significant differences at two grade levels. At Grade 2 boys scored significantly higher than girls on both subtests administered, Computations and Concepts. At Grade 6, boys again scored significantly higher than girls on two out of three subtests administered, Computations and Concepts. Statistical data appear in Appendix D.

d. Achievement Results - Howe vs. Parkwood

The achievement test results were examined for differences between the two IPI schools, and several significant differences were found. At Grade 2 children in Parkwood scored significantly higher than children at Howe on one of the two subtests administered, the Concepts subtest. At Grade 4 children at Howe scored significantly higher on the Computations subtest and scored higher to a degree that approached statistical significance on the Concepts subtest. This represents two out of three subtests at Grade 4. At Grade 6 children at Howe again scored significantly higher on the Computations subtest, which represents one out of three subtests administered at that grade level. In general, performance appeared somewhat better at Howe at the middle and upper grade levels and better at Parkwood at the lower grade levels. Statistical data appear in Appendix E.

e. Summary of Standardized Achievement Test Data

Table 4 summarizes the number of math subtests on which the IPI children at each grade level scored significantly

higher than Control children. The table shows the results as reported in each of the three years of program operation.

Table 4

Number* of Math Subtests on Which IPI Children Scored Higher Than Control Children at Each Grade Level in Each Year of Operation

Year 1		Year 2		Year 3	
Grade	Math Subtests	Grade	Math Subtests	Grade	Math Subtests
1	0/1	1	1/1	2	0/2
2	0/2	3	0/2	4	1/3
4	0/3	5	0/3	6	2/3
5	1/3				
6	0/3				

*Denominators refer to the number of subtests administered, and numerators represent the number of subtests on which IPI children scored higher

The table shows that over the three years of operation of the program there has been some improvement in the number of instances in which IPI children scored higher than Control children on the math subtests administered. No strong pattern of superiority in IPI schools has emerged, however. Considering the achievement results in Year 3 in light of the results in the previous two years of operation, it may be concluded that some, although limited, progress has been made toward attaining the achievement objective of the IPI Program.

f. Achievement in the IPI Math Continuum

In order to determine whether children were making progress within the IPI continuum, data were gathered to show the percentage of children working at various skill levels within the 13 areas of mathematics outlined by the program. These data were gathered from a sample of children at each of the grade levels 1 through 6 in September, 1971 and again in May of 1972. For the purposes of this evaluation report, the data from the two schools were combined.

Figure 3 shows the percentage of children working at different skill levels in September and in May of Year 3. The figure shows that the percentage of children working

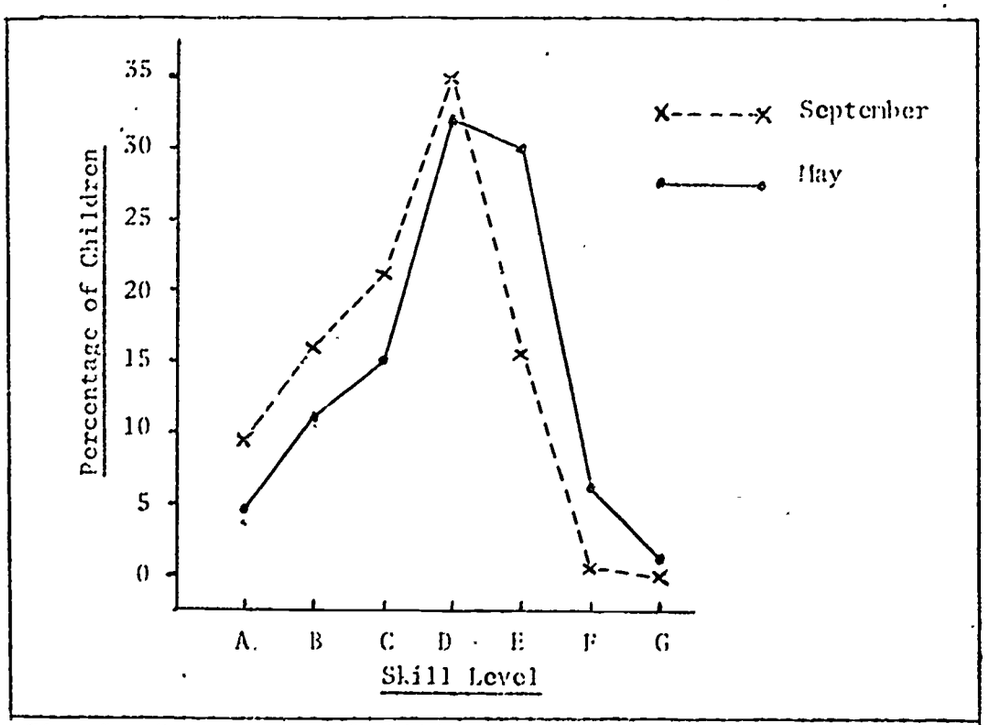

Fig. 3. Percentages of children working in IPI skill levels A through G in September and in May of Year 3.

at the lower skill levels (ABCD) decreased from September through May, while the percentage working in the higher range of skill levels (EFG) increased. These data indicate progress from the lower to the higher skill level range during the course of the year.

Figure 3 also shows that the percentage of children working in the highest skill level range (G) increased very little from September to May. In order to determine if the low percentage of children placing at the upper skill level was due to lack of progress by children in the upper grade levels, the data were analyzed by grade level. Figure 4 shows the percentage of children working at various skill levels for Grades 1 and 2 combined, Grades 3 and 4 combined, and Grades 5 and 6 combined. Again the percentages obtained in both

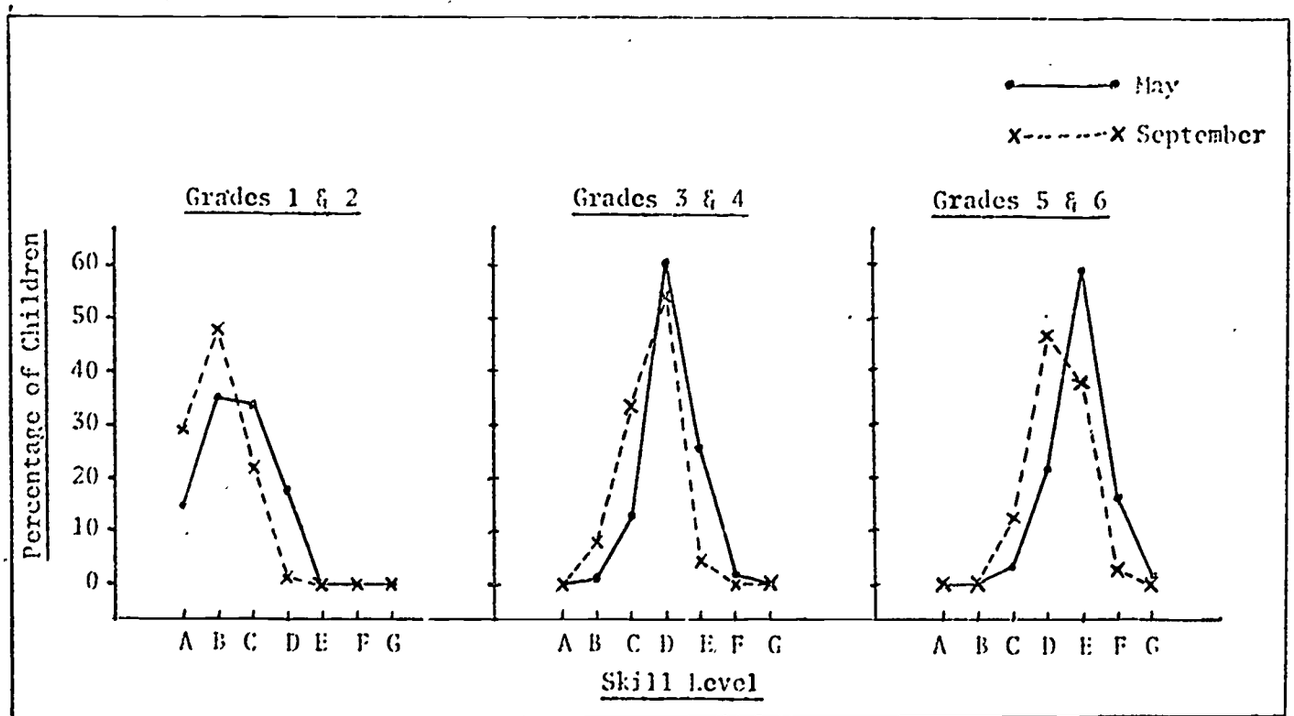

Fig. 4. Percentages of children working in IPI skill levels A through G in September and May of Year 3 for Grades 1 and 2, Grades 3 and 4, and Grades 5 and 6.

September and May are presented. The figure shows that for each of the three groups of children the median of the distribution increased by approximately one skill level across the 13 areas of mathematics from September to May, indicating that the children at all grade levels were progressing at approximately the same rate. The reason that more children did not place in the upper skill level appears to be that, given the observed rate of progress, few children were within striking range of these skill levels.

EVALUATION QUESTION 4

WHAT HAS BEEN THE IMPACT ON ACHIEVEMENT LEVELS IN MATH OVER THE THREE YEARS OF OPERATION OF THE PROGRAM?

1. General Procedures

To determine the impact on reading and math achievement over the full three years of operation of the IPI Program, two types of analysis were conducted. The first was a longitudinal analysis in which the achievement performance of a group of children was monitored throughout the three years of operation. The second analysis was cross sectional and involved monitoring changes in the achievement levels demonstrated at certain grade levels across the three years of program operation.

2. Longitudinal Analysis of Achievement

a. Longitudinal Data Collection

The schedule of achievement test administration over the three years of operation of the IPI Program allowed an examination of the performance of IPI and Control children on a longitudinal basis. The number

of children for whom data are available over the three year period and the grades in which they were tested are presented in Table 5. (The form and level of the tests administered at each grade in each year are available in Appendix F.)

Table 5

Sources of Longitudinal Samples Over Three Years of Operation of the IPI Program

Longitudinal Sample	Grades Tested			Longitudinal Sample Size	
	Year 1	Year 2	Year 3	IPI	Control
Gr. 1-2	---	Gr. 1	Gr. 2	31	68
Gr. 2-3-4	Gr. 2	Gr. 3	Gr. 4	35	62
Gr. 4-5-6	Gr. 4	Gr. 5	Gr. 6	72	127

Table 5 shows that the schedule of achievement testing generated three longitudinal samples. Two years of data are available on children who began the program in Grade 1, Year 2 and completed Grade 2, Year 3 (1-2 longitudinal sample). Three years of data are available on children who began the program in Grade 2, Year 1 and who completed Grade 4, Year 3 (2-3-4 longitudinal sample). Three years of data are also available on children who began the program in Grade 4, Year 1 and completed Grade 6, Year 3 (4-5-6 longitudinal sample).

The longitudinal analysis was performed on the achievement data of only those children who were enrolled in the IPI and Control schools during the entire longitudinal time frame. Children who entered or left the schools during the longitudinal time frame were not included in the analysis.

An important point to be remembered in interpreting the longitudinal analysis is that the combination of high pupil mobility rates and imperfect data retrieval methods resulted in a decrease in the size of the longitudinal sample each year. Consequently, the number of children in the longitudinal samples after three years represented only about one third of the children who were originally enrolled in the schools. The rest of the original enrollees had either left the schools or did not have complete achievement data. The sample that remained represented the most stable elements of the pupil population in terms of mobility and was in that sense a biased sample. The possible effect of this bias on achievement results is unknown.

b. Longitudinal Analysis - Results

Multivariate analysis of covariance was performed on the mean Year 3 post math scores of IPI and Control children in the three longitudinal samples. PLR scores and math pre scores from Year 1 were used as covariates for the Grade 2-3-4 and Grade 4-5-6 samples. Metropolitan Reading Readiness Test scores were used as covariates for the Grade 1-2 sample. Table 6 shows the mean post raw scores adjusted for unequal N and the effects of the covariates. Full statistical data are available in Appendix G.)

Table 6

Post Mean Math Scores for IPI and Control Children in the Grade 1-2, Grade 2-3-4 and Grade 4-5-6 Longitudinal Samples

		MEAN POST MATH SCORES					
Sample	Group	Computations	Superior Group	Concepts	Superior Group	Applications	Superior Group
Gr. 1-2	IPI	21.23	None	18.13	None	---	---
	Control	22.38		18.42		---	---
Gr. 2-3-4	IPI	14.01	Control*	14.73	None	11.55	None
	Control	17.47		15.36		10.70	
Gr. 4-5-6	IPI	11.76	Control**	11.65	Control***	14.47	None
	Control	17.98		12.41		13.83	

*$p < .0008$
**$p < .0001$
***$p < .085$

Table 6 shows that the only significant differences in any of the three longitudinal samples were in favor of the Control schools. In both the Grade 2-3-4 and Grade 4-5-6 samples, Control children scored significantly higher than IPI children on the Computations subtest. In the Grade 4-5-6 sample, Control children scored higher on the Concepts subtest to a degree that approached statistical significance.

To show the math performance of the longitudinal samples in relation to the grade level norms, the average raw scores of the IPI and Control children were transformed into grade equivalent scores. These are presented for each year of program operation in Figure 5.

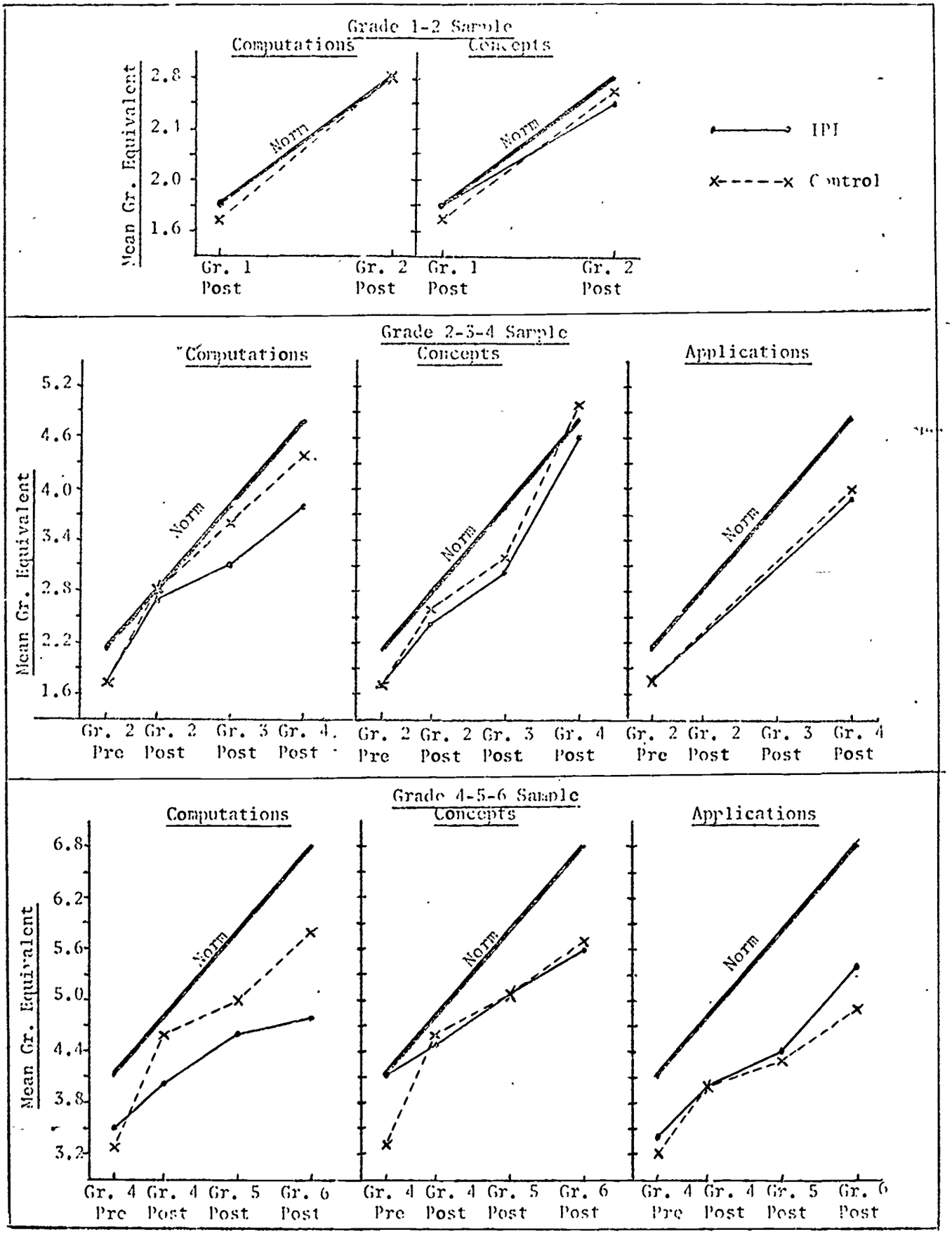

Fig. 5. Mean Math Grade Equivalent Scores for Three Longitudinal Samples of IPI and Control Children.

Figure 5 shows that at the end of the first year in the program the children in the Grade 1-2 sample in IPI schools were performing at grade level, but that at the end of two years, they had begun to fall behind the norm on the Concepts subtest. (The line representing the scores for the IPI children is not readily apparent in the graph showing performance on the Computations subtest, because it exactly coincides with the norm line.)

In the Grade 2-3-4 sample, children in the IPI schools fell progressively further behind the norms on the Computations and Applications subtests until by the end of three years in the program they were approximately one year below the norms. On the Computations subtest, the loss was greater than that demonstrated in Control schools. On the Concepts subtest, however, IPI children lost ground relative to the norms during the first two years but made substantial progress in closing the gap in the third year. By the end of three years in the program they were only two months below the norm on the Concepts subtest. Control children, who had started at the same level as the IPI children, were meanwhile performing two months above the norm at the end of three years. In the Grade 4-5-6 sample, a cumulative performance deficit was evident throughout the three years of operation for both IPI and Control children. On the Computations subtest the deficit was much greater for IPI children than for Control children, with the IPI children performing two years below the norm after three years in the program. On the Concepts subtest,

IPI children began the program achieving at the norm level, but fell a year behind the norms after three years in the program to a level comparable with that of Control children. On the Applications subtest the deficit was greater for Control children, but the IPI children had still fallen approximately a year and a half below the norms at the end of three years.

c. Summary of Longitudinal Analysis

Comparison of the math achievement results of children who had received the services of the IPI program continuously for two or three year periods with that of children enrolled in Control schools for the same period showed that where significant differences occurred, they were in favor of Control children. In most cases, children receiving IPI services for two or three years fell progressively further below the grade level norms over the course of time.

It should be noted that the longitudinal samples at the end of Year 3 are not exactly the same groups of children as the samples at the end of Year 2 because of pupil mobility. As mentioned earlier, with each succeeding year, the group of children in both IPI and Control schools who have been there since the program began gets smaller and more select. Restricting the longitudinal sample to geographically stable students may introduce a bias, the extent and effect of which is not known.

3. Cross Sectional Analysis of Achievement

a. Cross Sectional Data Collection

The cross sectional analysis of achievement attempts to answer the following type of question: After three years of operation of the IPI Program, how are children in Grade 3 performing as compared with the performance at Grade 3 when the program began? Cross sectional grade level comparisons of this sort were made using the results of the reading and arithmetic subtests of the Comprehensive Tests of Basic Skills (CTBS) obtained through the city-wide testing program. Mean scores on the subtests were obtained for the IPI schools, the Control schools and for all 30 Title I target schools. The test results from the 1969-70, the 1970-71 and the 1971-72 school years (Years 1, 2 and 3 of the IPI Program) were examined at Grades 3, 5 and 6. The use of these years and grades was determined by the city-wide testing schedule. Baseline data are not available for the years preceding the IPI Program because the CTBS was not used in the city-wide testing program until the first year of IPI.

b. Cross Sectional Data Analysis

Figure 6 shows the average math performance levels of children in Grades 3, 5 and 6 in the IPI schools, Control schools and the Title I target schools during each of the three years of operation of the IPI Program. Because test dates sometimes varied, the data are presented in the form of grade equivalent months deviation from the norm, rather

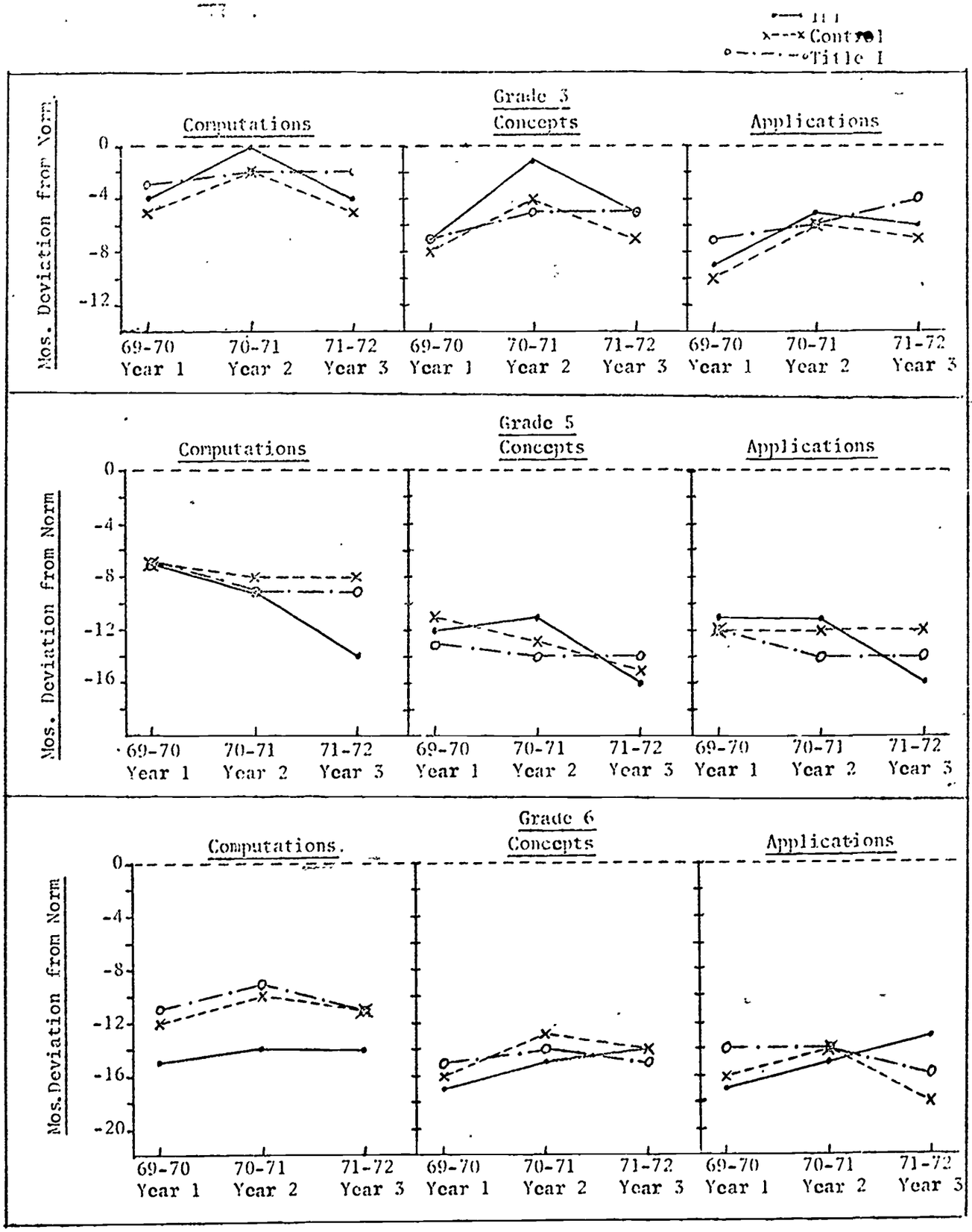

Fig. 6. Performance Levels in Math at Three Grade Levels Across Three Years for IPI, Control and 30 Title I Target Schools.

than grade equivalent scores. Due to the lack of statistical control on these data, apparent differences in the performance level between groups at a given point in time should not be interpreted too strictly. What is of interest is the pattern of performance changes by each group across the three years.

The figure shows that at Grade 3 the changes in math performance over the three years of the IPI Program were generally parallel for IPI and Control schools. Performance rose in the IPI schools from Year 1 to Year 2, but the same pattern was observed in the Control schools and the other Title I target schools. In Year 3 performance levels fell again in both IPI and Control schools, while it continued to rise or remained the same in the Title I target schools. The pattern of performance changes was similar on all three math subtests.

In Grade 5, there was again little difference between the IPI schools, the Control schools and the other Title I target schools in the pattern of performance changes between Year 1 and Year 2. In Year 3, however, the performance levels in the IPI schools dropped precipitously, although there was little change in the performance levels in the Control or other Title I target schools. The pattern of performance level changes was consistent across all three subtests administered.

In Grade 6, the pattern of change in performance levels was similar among all three groups of schools on the Computations subtest. On the Concepts and Applications subtests however, there was some difference between the changes

that occurred in the IPI schools and the changes that occurred in the Control and other Title I schools. Between Year 1 and Year 2 the changes in performance were similar for the three groups, a slight rise in performance levels. In Year 3, however, performance levels dropped in the Control and other Title I schools, but continued to rise in the IPI schools, In Year 3 performance in the IPI schools was four months higher than in Year 1 on both the Concepts and Applications subtests. In the Control and other Title I schools performance levels in Year 3 were generally the same or lower than in Year 1, especially on the Applications subtests.

c. Summary of Cross Sectional Analysis

In examining changes in math performance at three grade levels across the three years of operation of the IPI Program, any such changes must be interpreted in light of the changes occurring in other schools. In this cross sectional look at achievement, there was some evidence of impact of the IPI Program at Grade 6: on two subtests a pattern of slowly improving performance in the IPI schools took place in the face of generally declining performance in other schools. The pattern was by no means conclusive, however, and several more years of data would be required before any definitive statements could be made. In Grade 3 the changes in performance levels in IPI schools were generally paralleled by similar changes in the Control and other Title 1 target schools. At Grade 5 the performance changes in the IPI schools compared unfavorably

with the changes in the other schools. In general, the cross sectional analysis yielded little evidence that the IPI Program has had a solid impact on achievement levels in the schools served by the program.

V. CONCLUSIONS AND RECOMMENDATIONS

A. Discussion of Results

The results of the third year of operation of the IPI Program have shown that (as far as can be determined without a detailed analysis of specific diagnostic and prescription writing procedures) the IPI Program continues to be implemented according to plan. The question that remains is the effectiveness of the program in raising the math achievement levels of the children served. The pre-post analysis of progress through the skill levels within the IPI continuum showed that the children are moving along the continuum, but because of the lack of norms against which to compare their progress, it is difficult to determine whether they are progressing at a "normal" rate. The evaluator must, therefore, rely on information from standardized achievement tests to draw conclusions about the impact of the program on math performance.

The achievement data from the third year of operation showed that the children in IPI schools fared somewhat better in comparisons with children in Control schools than they did in previous years. The results, however, did not indicate a strong impact, except at Grade 6. Similarly, cross sectional

analyses of changes in performance levels at given grade levels across the three years of operation showed little evidence of improvement except at Grade 6. Longitudinal analyses of the performance of only those children who had been in the program continuously for two or three years showed no evidence at all of superior performance by IPI children.

As discussed in the evaluation report for Year 2 of operation, there are several possible reasons why a greater impact on achievement has not been evident. The first possibility is, of course, that the IPI approach simply does not work any better than the traditional methods of math instruction. Other possibilities exist, however. It is possible that the standardized achievement results do not adequately measure what is taught in the IPI continuum, (a possibility that raises questions as to what is taught in the IPI continuum). One way of investigating this thesis is to analyze the standardized tests item by item to determine whether the questions reflect the skills and concepts covered in IPI. The team of consultant teachers in Stephen Howe school was assigned this task, and their analysis of the arithmetic subtests of the Stanford Primary I, Stanford Primary II, Stanford Intermediate I and Stanford Intermediate II achievement tests showed that an average of 94% of the items on these tests were covered in the IPI continuum. The range of correspondence between the Stanford tests and the IPI continuum was from 82% to 100% for nine arithmetic subtests. It would appear that the standardized tests do measure the same things that are taught in the IPI continuum.

It is possible, of course, even though the standardized tests measure basically the same skills as are taught in IPI, that the language of the tests is sufficiently different from the language used in IPI to impair performance on the tests. For example, one question in the Stanford Intermediate 1 Arithmetic Concepts makes use of the terms multiplier, multiplicand and minuend, but these terms are not used at all in IPI. To investigate this potential problem, the consultant teachers in the IPI schools constructed achievement tests directly from the IPI materials. Tests were constructed for Grades 3 and 5, and items were selected that, in the judgment of the IPI staff, reflected the instructional activities in which all or most of the pupils had participated. Since the test items reflected what most of the children had been taught in the IPI continuum, a mastery level performance was expected. It was anticipated that each item would be answered correctly by at least 75% of the children. The results showed, however, that no more than 36% of the items on any of these tests were correctly answered by 75% or more of the children. In fact no more than 57% of the items on any of the tests were correctly answered by more than 60% of the children. These results are not conclusive, because these IPI tests are as yet in unrefined form, but they do suggest that lack of evidence of improvement in math achievement may not be due to the use of standardized tests.

Another possibility for explaining the apparent lack of progress in raising achievement levels is that the IPI

system teaches the math skills that the Cleveland Schools wants taught, but that it teaches them in a different sequence from that normally followed. That is, by the time he leaves elementary school, a child in IPI may have covered the same concepts and skills as a child not in IPI (hopefully with greater mastery), but at any given point in time during their elementary school careers, they may be working on entirely different skills. If this is the case, the only valid assessment of IPI possible is a longitudinal study in which achievement levels of IPI and Control children are compared only after the IPI children have spent all or most of their elementary school years working in the IPI system. One problem with this approach, however, is that the percentage of children remaining in the schools that long is relatively small. Less than one third of the children present in the IPI schools during Year 1 were still in the schools by the end of Year 3. Therefore, the whole question of whether the IPI Program might have an impact on a cvement if children participated for their entire elementary school tenure may be academic.

Finally, despite the indications that the program guidelines have been faithfully followed by the IPI staff, over the course of three years it is possible that deviations from the precise testing and prescription criteria established by the program's designers have crept into the operation. If so, these deviations may be reflected in lack of improvement in math achievement. It was recommended in the evaluation report

for Year 2 that another evaluation of operations accuracy be conducted by Research for Better Schools, but that organization has reported that they no longer perform that service. Self-monitoring instruments have been developed, however, and their use in the Cleveland IPI schools may yield valuable information as to the accuracy with which the IPI procedures are being followed.

In the long run, the decision about the future of the IPI Program will have to balance program costs against program benefits. There is some evidence that achievement levels in math in the IPI schools may be improving slightly at certain grade levels. But is is also clear that over the course of three years the cost of the program has risen 74% to $210 per pupil. The question becomes whether the present or future gains in achievement are great enough to justify the expense; or whether more efficient and less costly avenues to the same ends can be found.

At any rate, the results of field tests of experimental programs in the schools must be interpreted in light of the number of influences over which the evaluator has no control; experimental treatment populations change, control populations change, samples may become biased through non-random attrition. All of these factors may cause changes in performance that have nothing to do with the experimental treatment under investigation. Evaluation must, therefore, take a long and cautious view of results and data must be examined over a period of several years before administrators can state with some degree of confidence that a given program is or is not effective.

B. Recommendations

1. It is recommended that the IPI Program be continued in the same two schools in which it presently operates for the next one or two years until sufficient data are available for an accurate appraisal of long term program effects.

2. It is recommended that a decision on expansion of the IPI Program to other schools be deferred until the results of a long term evaluation are available.

3. It is recommended that the development and refinement of achievement tests constructed from the IPI materials proceed.

4. It is recommended that self-monitoring procedures be instituted in each IPI school, using the instruments developed for this purpose by RBS.

A P P E N D I X

APPENDIX A

Descriptive Data on Project Schools

	Parkwood	Stephen Howe
Poverty Rate	46%	38%
Mobility Rate	51%	89%
Enrollment in IPI Program by Grade:		
EMR	18	0
1	69	81
2	91	77
3	67	80
4	52	75
5	61	56
6	79	77
Total	437	446

APPENDIX B

MATCHED CHARACTERISTICS OF IPI SCHOOLS AND CONTROL SCHOOLS IN 1969

School	% Poverty	% Mobility	% Attendance	Grade	MEDIAN ARITHMETIC GRADE EQUIVALENTS[1] Computation	Concepts	Application	Mean[2] PLR	% Negro
Parkwood	36	23	94.6	6	5.6	5.4	4.9	94	100
				5	4.4	5.2	4.2	101	
				3	3.7	3.7		95.3	
Hazeldell	29	20	94.3	6	5.3	5.4	4.9	96	100
				5	4.1	4.9	4.0	99	
				3	3.4	2.8		93.2	
S.E. Howe	31	32	94.4	6	6.2	6.1	4.9	99	100
				5	5.6	5.4	4.6	98	
				3	3.7	3.3		103.1	
Hough	40	35	90.6	6	6.0	5.9	5.1	98	100
				5	5.2	5.6	4.4	97	
				3	3.5	2.9		88.6	

[1]Results of Stanford Primary II and Intermediate II Achievement Tests. City-Wide Testing, Spring, 1969.

[2]Probable Learning Rate. Score on Kuhlman Anderson Test.

APPENDIX C

Math Achievement - IPI vs. Control

Adjusted Mean Pre and Post Math Achievement Test Scores
for IPI and Control Schools at Grade 2 with
Results of Multivariate Analysis of Covariance

Subtest	Mean Pre Scores (Covariates)		Mean Post Scores		Least Square Estimate	Univarite F	df
	IPI	Control	IPI	Control			
Arithmetic	36.68	34.77	---	---	---	---	---
Arithmetic Computations	---	---	21.22	21.81	-.592	.03	1 & 209
Arithmetic Concepts	---	---	16.29	16.90	-.608	.28	1 & 209
PLR	105.3	114.4		Multivariate F = .14			2 & 208

NOTE: Analysis was performed on post test scores with pre test and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects. Stanford Primary I used as pre test and Stanford Primary II used as post test.

The Stanford Primary I given as the pre test in Grade 2 has only one arithmetic subtest. The scores from this subtest were used as the pre-scores for the Computations and Concepts subtests in the Primary II, which was given as the post test.

APPENDIX C Continued

Math Achievement - IPI vs. Control

Adjusted Mean Pre and Post Math Achievement Test Scores for IPI and Control Schools at Grade 4 with Results of Multivariate Analysis of Covariance

Subtest	Mean Pre Scores (Covariates)		Mean Post Scores		Least Square Estimate	Univariate F	df
	IPI	Control	IPI	Control			
Arithmetic Computations	8.58	10.04	12.84	16.64	-3.80	31.92*	1 & 206
Arithmetic Concepts	8.77	9.43	13.44	14.61	-1.17	3.72**	1 & 206
Arithmetic Applications	8.21	8.96	11.23	10.38	.85	2.85***	1 & 206
PLR	99.96	95.20		Multivariate F = 18.67*			3 & 204

*$p < .0001$
**$p < .055$
***$p < .093$

NOTE: Analysis was performed on post test scores with pre test and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects. Alternate forms of Stanford Intermediate 1 used as pre and post tests.

APPENDIX C Continued

Math Achievement - IPI vs. Control

Adjusted Mean Pre and Post Math Achievement Test Scores for IPI and Control Schools at Grade 6 with Results of Multivariate Analysis of Covariance

Subtest	Mean Pre Scores (Covariates)		Mean Post Scores		Least Square Estimate	Univariate F	df
	IPI	Control	IPI	Control			
Arithmetic Computations	9.95	11.38	13.70	16.29	- 2.59	26.39*	1 and 323
Arithmetic Concepts	9.00	9.76	12.27	11.38	.89	5.59**	1 and 323
Arithmetic Applications	12.38	11.30	15.24	13.22	2.03	17.63*	1 and 323
PLR	96.32	98.15		Multivariate F = 21.58*			3 and 321

*$p < .0001$
**$p < .019$

NOTE: Analysis was performed on post test scores with pre test and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects.
Alternate forms of Stanford Intermediate II used as pre and post tests.

APPENDIX C Continued

Achievement - Sex x Treatment Interaction

Adjusted Mean Pre and Post Achievement Test Scores for IPI and Control Boys and Girls at Grade 2 with Results of Multivariate Analysis of Covariance

Subtest	Mean Pre Scores (Covariates)				Mean Post Scores				Interaction Least Square	Interaction Univariate F	df
	IPI		Control		IPI		Control				
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls			
Arithmetic	36.23	37.14	33.50	36.93	---	---	---	---	- 4.49	4.93*	1 and 209
Arithmetic Computations	---	---	---	---	21.05	21.38	23.89	19.73	- 2.44	2.31	1 and 209
Arithmetic Concepts	---	---	---	---	17.10	15.49	18.93	14.88			
PLR	103.8	106.8	112.9	116.0	---	---	---	---	Multivariate F = 3.25**		2 and 208

*$p < .028$
**$p < .042$

NOTE: Analysis was performed on post test scores with pre test and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects.

APPENDIX D

Achievement - Boys vs. Girls

Adjusted Mean Pre and Post Achievement Test Scores for Boys and Girls at Grade 2 with Results of Multivariate Analysis of Covariance

Subtest	Mean Pre Scores (Covariates)		Mean Post Scores		Least Square Estimate	Univariate F	df
	Boys	Girls	Boys	Girls			
Arithmetic[a]	34.56	36.58	---	---	---	---	---
Arithmetic Computations	---	---	22.47	20.56	1.92	3.92**	1 & 209
Arithmetic Concepts	---	---	18.01	15.19	2.83	19.15*	1 & 209
PLR	108.3	111.4		Multivariate F = 10.55*			2 & 208

*$p < .0001$
**$p < .05$

NOTE: Analysis was performed on post test scores with pre test scores and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects. Stanford Primary I used as pre test and Stanford Primary II used as post test.

[a]The Stanford Primary I Achievement Test given as the pre test in Grade 2 has only one arithmetic subtest. The scores from this subtest were used as the pre scores for the Computations and Concepts subtests in the Stanford Primary II which was given as the post test.

APPENDIX D Continued

Achievement - Boys vs. Girls

Adjusted Mean Pre and Post Achievement Test Scores for Boys and Girls at Grade 4 with Results of Multivariate Analysis of Covariance

Subtest	Mean Pre Scores (Covariates)		Mean Post Scores		Least Square Estimate	Univariate F	df
	Boys	Girls	Boys	Girls			
Arithmetic Computations	8.36	10.25	14.68	14.80	- .11	.21	1 & 206
Arithmetic Concepts	9.15	9.04	14.29	13.77	.51	.69	1 & 206
Arithmetic Applications	8.58	8.58	11.25	10.36	.88	3.63	1 & 206
PLR	94.68	99.57		Multivariate F = 1.99*			3 & 204

*$p < .05$

NOTE: Analysis was performed on post test scores with pre test and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects.

APPENDIX D Continued

Achievement - Boys vs. Girls

Adjusted Mean Pre and Post Achievement Test Scores for Boys and Girls at Grade 6 with Results of Multivariate Analysis of Covariance

Subtest	Mean Pre Scores (Covariates)		Mean Post Scores		Least Square Estimate	Univariate F	df
	Boys	Girls	Boys	Girls			
Arithmetic Computations	10.01	11.32	14.33	15.65	- 1.32	5.37***	1 & 323
Arithmetic Concepts	9.69	9.08	12.18	11.47	.71	5.91**	1 & 323
Arithmetic Applications	11.69	12.00	14.49	13.97	.52	1.66	1 & 323
PLR	95.65	98.82		Multivariate F = 5.80*			3 & 321

*$p < .0008$
**$p < .016$
***$p < .022$

NOTE: Analysis was performed on post test scores with pre test and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects.

APPENDIX E

Achievement - Howe vs. Parkwood

Adjusted Mean Pre and Post Achievement Test Scores for Stephen Howe and Parkwood Schools at Grade 2 with Results of Multivariate Analysis of Covariance

Subtest	Mean Pre Scores (Covariates)		Mean Post Scores		Least Square Estimate	Univariate F	df
	Howe	Parkwood	Howe	Parkwood			
Arithmetic[a]	39.11	34.25	---	---	---	---	---
Arithmetic Computations	---	---	20.22	22.22	- 2.00	1.94	1 & 209
Arithmetic Concepts	---	---	14.23	18.36	- 4.13	17.89*	1 & 209
PLR	108.9	101.7		Multivariate F = 9.27**			2 & 208

*p < .0001
**p < .0002

NOTE: Analysis was performed on post test scores with pre test scores and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects. Stanford Primary I used as pre test and Stanford Primary II used as post test.

[a]The Stanford Primary 1 Achievement Test given as the pre test in Grade 2 has only one arithmetic subtest. The scores from this subtest were used as the pre scores for the Computations and Concepts subtests in the Stanford Primary II which was given as the post test.

APPENDIX L Continued

Achievement - Howe vs. Parkwood

Adjusted Mean Pre and Post Achievement Test Scores for Stephen Howe and Parkwood Schools at Grade 4 with Results of Multivariate Analysis of Covariance

Subtest	Mean Pre Scores (Covariates)		Mean Post Scores		Least Square Estimate	Univariate F	df
	Howe	Parkwood	Howe	Parkwood			
Arithmetic Computations	9.07	8.09	14.58	11.10	3.48	10.13*	1 & 206
Arithmetic Concepts	8.69	8.84	14.25	12.64	1.61	3.56***	1 & 206
Arithmetic Applications	8.89	7.52	11.27	11.18	.09	.27	1 & 206
PLR	98.93	99.13		Multivariate F = 4.02**			3 & 204

*$p < .002$
**$p < .009$
***$p < .061$

NOTE: Analysis was performed on post test scores with pre test and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects.

APPENDIX F Continued

Achievement - Howe vs. Parkwood

Adjusted Mean Pre and Post Achievement Test Scores for Stephen Howe and Parkwood Schools at Grade 6 with Results of Multivariate Analysis of Covariance

Subtest	Mean Pre Scores (Covariates)		Mean Post Scores		Least Square Estimate	Univariate F	df
	Howe	Parkwood	Howe	Parkwood			
Arithmetic Computations	11.88	8.02	14.88	12.52	2.35	10.18*	1 & 323
Arithmetic Concepts	9.93	8.07	12.27	12.27	-.001	.003	1 & 323
Arithmetic Applications	13.83	10.93	14.91	15.58	-.67	1.16	1 & 323
PLR	100.1	92.55		Multivariate F = 4.56**			3 & 321

*$p < .002$
**$p < .004$

NOTE: Analysis was performed on post test scores with pre test and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects.

APPENDIX F

Form and Level of Stanford Achievement Tests Administered In Year 1, Year 2 and Year 3 to Pupils in the Longitudinal Samples.

Level and Form of Test	Longitudinal Sample		Sample Size	
	Grade	Program Year	IPI	Control
Post - Primary I, W	1	2	31	68
Pre - Primary I, X Post - Primary II, W	2	3		
Pre - Primary I, W Post - Primary II, W	2	1	35	62
Pre - Primary II, X Post - Primary II, W	3	2		
Pre - Intermediate I, X Post - Intermediate I, W	4	3		
Pre - Intermediate I, W Post - Intermediate I, X	4	1	72	127
Pre - Intermediate II, X Post - Intermediate II, W	5	2		
Pre - Intermediate II, X Post - Intermediate II, W	6	3		

APPENDIX G

Achievement - IPI vs. Control Longitudinal Data

Adjusted Mean Post Achievement Test Scores for IPI and Control Children in the Longitudinal Samples with Results of Multivariate Analysis of Covariance

Grade 1 - 2

Subtest	Mean Post Scores		Least Square Estimate	Univariate F	df
	IPI	Control			
Arithmetic Computations	21.23	22.38	1.15	.79	1 and 90
Arithmetic Concepts	18.15	18.12	- .29	.003	1 and 90
Metropolitan Reading Readiness Test (Covariate)	72.05	73.46	---	---	---
			Multivariate F = .45		2 and 89

NOTE: Analysis was performed on Grade 1 post test scores with Metropolitan Reading Readiness Test score as the covariate. Covariate shown is adjusted for unequal N. Post test scores shown are adjusted for unequal N and the effect of the covariate.

APPENDIX G Continued

Achievement - IPI vs. Control Longitudinal Data

Adjusted Mean Pre and Post Achievement Test Scores for IPI and Control Children in the Longitudinal Samples with Results of Multivariate Analysis of Covariance

Grade 2-3-4

Subtest	Mean Pre Scores (Covariates)		Mean Post Scores		Least Square Estimate	Univariate F	df
	IPI	Control	IPI	Control			
Arithmetic[a]	30.00	33.46	---	---	--	---	---
Arithmetic Computation	---	---	14.01	17.47	- 3.47	12.22**	1 & 87
Arithmetic Concepts	---	---	14.73	15.36	- .63	.52	1 & 87
Arithmetic Applications	---	---	11.55	10.70	.85	1.23	1 & 87
PLR	101.0	109.1		Multivariate F = 7.42*			3 & 85

*$p < .0002$
**$p < .0008$

NOTE: Analysis was performed on Grade 4 post test scores with Grade 2 pre test and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects.

[a]The Stanford Primary 1 Achievement Test given as the pre test in Grade 2 has only one aritimetic subtest. The scores from this subtest were used as the pre scores for the Computations, Concepts and Applications subtests in the Stanford Intermediate I which was given as the post-test.

APPENDIX G Continued

Achievement - IPI vs. Control Longitudinal Data

Adjusted Mean Pre and Post Achievement Test Scores for IPI and Control Children in the Longitudinal Samples with Results of Multivariate Analysis of Covariance

Grade 4-5-6

Subtest	Mean Pre Scores (Covariates)		Mean Post Scores		Least Square Estimate	Univariate F	df
	IPI	Control	IPI	Control			
Arithmetic Computations	10.15	9.03	11.76	17.98	- 6.22	87.39*	1 & 187
Arithmetic Concepts	9.52	9.61	11.56	12.41	- .76	3.07**	1 & 187
Arithmetic Applications	7.87	6.83	14.47	13.88	.59	.51	1 & 187
PLR	99.24	92.85		Multivariate F = 34.30*			3 & 185

*$p < .0001$
**$p < .082$

NOTE: Analysis was performed on Grade 6 post test scores with Grade 4 pre test and PLR scores as covariates. Covariates shown are adjusted for unequal N. Post test scores shown are adjusted for unequal N and covariate effects.