DOCUMENT RESUME

ED 216 339 CS 006 690

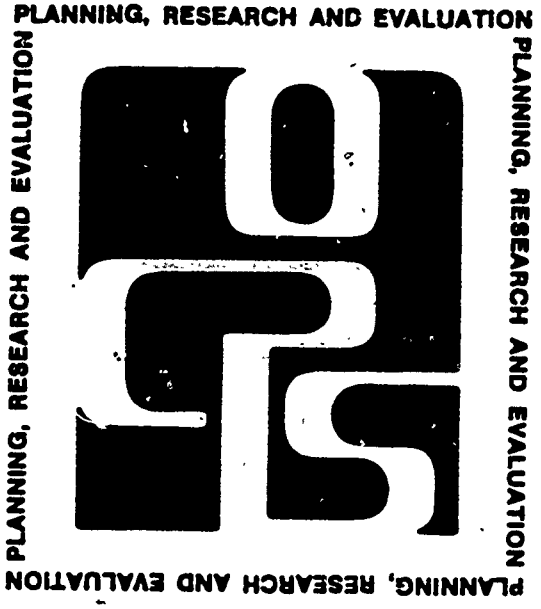
AUTHOR Kimball, George H.
TITLE The Effects of Utilization of PRI/RS on Reading Achievement in the Oklahoma City Public Schools. Evaluation of 1981-82.
INSTITUTION Oklahoma City Public Schools, OK. Dept. of Planning, Research, and Evaluation.
PUB DATE May 82
NOTE 59p.
JOURNAL CIT Journal of Research and Evaluation of the Oklahoma City Public Schools; v12 n1 May 1982

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Classroom Techniques; *Computer Managed Instruction; *Criterion Referenced Tests; Elementary Education; Evaluation Methods; Management Systems; *Reading Achievement; *Reading Instruction; *Reading Programs; *Reading Research; Reading Tests; Systems Approach; Validity
IDENTIFIERS Prescriptive Reading Inventory; *Skills Management System

ABSTRACT

A study was conducted to determine (1) the extent to which the degree of use of PRI/RS (the Prescriptive Reading Inventory Reading Systems, a criterion referenced management system developed by CTB/McGraw Hill) affected comparable groups of students on standardized achievement measures; (2) whether the degree of use differentially influenced students' achievement based on their initial ability levels; and (3) the extent to which differential use of PRI/RS and TRACER, a computer management system, affected reading acquisition measures, such as number of objectives mastered and retention of skills mastered. Subjects were all students in grades two and five in a large city school district. Analysis of data showed that the use of PRI/RS had a systematic and reliable effect on student reading achievement. The students whose teachers used PRI/RS and TRACER consistently and to a high degree performed significantly better on standardized achievement tests. It also appeared that the effect was distributed across ability levels. At all ability levels, the high-use individuals demonstrated superior reading achievement test scores. (FL)

JOURNAL OF RESEARCH AND EVALUATION
OF THE OKLAHOMA CITY PUBLIC SCHOOLS

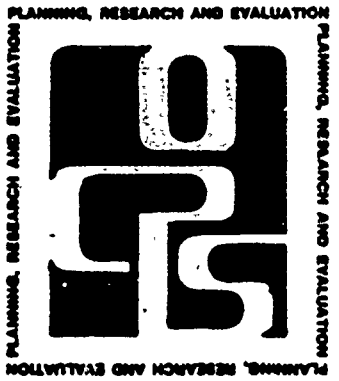


EVALUATION OF 1981-82 EFFECTS OF UTILIZATION OF PRI/RS

OKLAHOMA CITY PUBLIC SCHOOLS
THOMAS PAYZANT, SUPERINTENDENT
OKLAHOMA CITY, OKLAHOMA

VOLUME 12. . . .NUMBER 1
MAY 1982

THE EFFECTS OF UTILIZATION OF PRI/RS ON READING ACHIEVEMENT IN THE OKLAHOMA CITY PUBLIC SCHOOLS



By
George H. Kimball, Ph.D.
Research Associate

Planning, Research, and Evaluation Department
Oklahoma City Public Schools
Oklahoma City, Oklahoma

EXECUTIVE SUMMARY

In 1980-81, the Prescriptive Reading Inventory Reading Systems (PRI/RS), Criterion-referenced management system from CTB/McGraw Hill was implemented in the Oklahoma City Public Schools elementary grades (K-5). During the 1981-82 school year, PRI/RS was expanded to include the Middle School grades (6-8), and the Mathematics equivalent of PRI/RS (Diagnostic Mathematics Inventory or DMI) was instituted as well (in grades 6-8 only). The PRI/RS was selected over other commercially available criterion-referenced reading systems (as well as over the possiblity of custom designing a system unique to the Oklahoma City Public Schools) because it was felt that:

1. PRI/RS had a high degree of fit (97% match of the objectives) to the district's established reading/language arts objectives.
2. PRI/RS has been successfully used in other large school districts.
3. As a management device, it would help facilitate the coordination of instruction between the regular classroom teacher, Special Education and Title I teachers, and also facilitate the continuity of instruction between all teachers and schools within the district.

4. The publishers had substantiated the validity and reliability of its criterion-referenced testing development procedures.

5. The criterion-referenced mastery tests were readily available and compatible for use with a computer management system (TRACER) already in use in the Oklahoma City Public Schools.

6. The instruction, and support materials were well planned, organized, skill defined, and provided illustrated procedures for reading skill instruction and management.

Purpose of the Study

The purpose of this analysis is multifaceted. The questions to be addressed include: (1) "To what extent does the degree of utilization of PRI/RS affect comparable groups of students on standardized achievement measures?" (2) "Does degree of utilization differentially influence students' achievement based on their initial ability levels?" (i.e., do students benefit more or less depending on their ability levels?) and (3) "To what extent does differential utilization of PRI/RS and TRACER affect reading acquisition measures such as number of objectives mastered on the PRI/RS diagnostic test, and retention of skills mastered.

Effects of Utilization on Standardized Achievement Scores

For the analysis relevant to the first question, "To what extent does degree of utilization affect comparable groups of students on standardized achievement measures?", groups of

students in the second and fifth grades were utilized. Since random assignment of students to treatment conditions was not possible, a stratified random selection procedure was utilized within the established treatment groups, so that students of all ability levels (from chance level responders to 99th percentile individuals) were equally represented in both groups. The level of PRI/RS utilization was determined by the consistency of TRACER utilization and the total number of student mastery tests submitted by their classroom teachers. Two groups were established: (1) high utilization and (2) low utilization.

Caution needs to be observed in interpreting the designation of "high" and "low" utilization groups. The differentiation of high and low utilization individuals does not infer that no effective teaching behaviors were occuring in the low utilization group. It is assumed that low utilization teachers were: (1) proceeding with classroom behaviors which were typical prior to implementation of PRI/RS; and (2) were continuing to utilize management techniques and materials with which they were comfortable and familiar. It is also assumed that the students benefited from these teaching behaviors and techniques. Examination of the California Achievement Test (CAT) scores showed that gains in reading were demonstrated by students in both groups. Further, the results of this study do not rely on a comparison vs. a no-treatment control group. Rather, the intent is to interpret the degree to which the high utilization of the regimen

of the PRI/RS criterion-referenced testing program, and the individualized record of student growth provided by TRACER, produces superior student achievement gain relative to "base line" data (i.e., the achievement of students taught under a traditional, non-criterion-referenced approach).

For both Grades 2 and 5, the Analysis of Covariance (ANCOVA) results indicated that the high utilization of PRI/RS was influential in producing a higher raw score (group mean) on the Total Reading Composite of the California Achievement Test.

Effects of Utilization for Students' of Different Ability Levels

To address question number two, "Does degree of utilization differentially influence students' achievement based on their initial ability levels?", the students were partitioned into three categories (according to their initial achievement levels). This grouping was based on their spring 1980 Total Reading Composite of the CAT (national percentiles). The groups were: (1) Low Achievers (chance level through 40% percentile), (2) Average Achievers (41st through 80th percentile), and (3) High Achievers (81st through 99th percentile). This, in effect, produced an Analysis of Variance design of initial achievement levels by utilization (three levels of achievement, and two levels of treatment).

The results of the ANOVA showed that the between group differences (high vs. low utilization) were consistent across all ability levels. There was a significant difference between the groups at each ability level, which suggests that the utilization of the PRI/RS approach benefits students of all ability levels, when they are compared to their counterparts taught under a non-criterion-referenced approach.

The Effect of Utilization on the Number of Reading Objectives Mastered

In addition to determining the influence of the criterion-referenced approach on standardized achievement scores (CAT Total Reading Composite), the third question to be addressed was: "To what extent does differential utilization of PRI/RS affect reading reading acquisition measures such as the total number of reading objectives mastered on the PRI/RS diagnostic test, as well as the degree of retention of the reading objectives mastered.

For this analysis, the same groups of 1980-81 fifth grade students were utilized, as these students were the only group in the district to have taken the PRI/RS diagnostic test (12-5-80) as Elementary School students, and again this year (9-23-81) as Middle School students. This comparison provided

an independent pre-post measure of reading acquisition, on such variables as: (1) the number of reading objectives mastered (from diagnostic to diagnostic test), (2) the number of reading objectives retained (from mastery testing to diagnostic testing and diagnostic to diagnostic), (3) and the proportion of objectives retained and lost (mastery demonstrated on pre but not on post).

If the group differences based on degree of utilization apparent in the analysis on CAT results is reflecting a true difference in reading achievement, the effects should replicate in this additional analysis, based on the same 5th grade groups and degree of utilization of the system. As an example, the high utilization group would be expected to have a higher number of individual mastery tests submitted, and a higher total number of objectives mastered (prior to the post diagnostic test) by the very definition of the group. Yet, it is possible that the lower utilization teachers accomplished the amount of reading objectives mastered, yet simply did not utilize the PRI/RS mastery tests or TRACER to quantify the gains. If there is a significant learning difference between groups, they should also differ on the second PRI/RS diagnostic test (assuming "matched" levels of the diagnostic test were taken) on the number of objectives mastered.

Results

A multiple regression procedure was utilized to identify what variables had significant influence on the acquisition and retention of the various reading objectives. The analysis showed that the only variable having significant influence on both acquisition and retention, was the degree of utilization of PRI/RS. On the first acquisition measure, the total number of reading objectives mastered on the post-diagnostic test, the number of objectives mastered on the pre-test (their starting position) and the number of days the student was absent from school, also were indicated as being powerful influences on the total number of reading objectives mastered.

For the increase in number of reading objectives mastered between pre and post-diagnostic testing (which does not include those reading objectives duplicated in mastery on both tests, or those not retained between pre and post testing), the degree of utilization and the amount of days absent from school were the most influential factors in producing gains.

For retention measures (the number of reading objectives retained or duplicated between pre and post-diagnostic tests, as well as the number lost), the degree of utilization of PRI/RS was again most influential. For the percentage of duplication of objectives mastered (pre-to-post), the number of objectives mastered on the pre-test was also a significant influence.

Analysis of differences between the high and low utilization groups revealed that, in each case, the high utilization group performed significantly better than their matched counterparts in the low utilization group. The single exception was on a measure of the ratio of mastery tests submitted to the subsequent demonstration of mastery on the diagnostic test. In that comparison, the groups were not significantly different.

Summary and Conclusions

To summarize, and put into perspective the overall results of this study, it is necessary to remember that these designs can best be categorized as "quasi-experimental" designs. Without randomized assignment of subjects to treatment and control groups (with each subject having equal opportunity to be assigned to either), and without the added precision of a no-treatment control group for comparisons (which is possible in a laboratory situation), the drawing of causal inferences about treatment effects is tenuous.

However, given the a priori knowledge that in a quasi-experimental design, the groups cannot be considered equivalent, every possible effort was made to equate (statistically and conceptually) the groups, so that perceived differences can be attributed to treatment effects. This involved attending to

relevant variables which could not be controlled, and providing an index as to the effect these uncontrolled variables may have had on the demonstrated between-group differences (see the full technical report, for the description and analyses germane to those uncontrolled variables' influence).

It is further assumed that in studies such as this, where realistic knowledge of the effects of an instituted treatment has priority for decision-makers, over the ability to delineate single theoretical constructs (which may be presumed to be the casual agents in the learning process), scientific analyses within the constraints of the data are appropriate.

Within the contraints of quasi-experimental design and analysis, it appears that the utilization of PRI/RS has a systematic and reliable effect on reading achievement. Those 2nd and 5th grade students whose teachers utilized PRI/RS and TRACER consistently and to a high degree performed significantly better on the Total Reading Composite of the California Achievement Test. It also appears that the effect is distributed across ability levels. At all ability levels, the high utilization individuals demonstrated superior reading achievement results.

The effects of utilization on standardized achievement scores were consistent between Grade 2 and Grade 5. Further, the effects on Grade 5 achievement, appeared to replicate when other

independent measures of achievement (acquisition) and retention were utilized. The degree of utilization appears to differentially affect the number of objectives mastered on the PRI/RS diagnostic test, the number of increase of reading objectives, the number of objectives "lost" to mastery between pre-to-post diagnostic tests.

In summary, even though the utilization of PRI/RS cannot be proved with these data to be the "only" casual agent producing these results, it certainly appears to be influential in the hypothesized direction, and is at least partially causal in conjunction with other uncontrolled or unidentified variables.

A final word of caution about the limited scope of this analysis. This research is an attempt to quantify and identify the effect utilization of PRI/RS may have on Reading Achievement, as the PRI/RS becomes more functionally integrated and utilized system-wide in the Oklahoma City Public Schools. It is not, nor was it intended to be, an encompassing "Evaluation" of the PRI/RS system in the OKC Public Schools. No data are available to identify why some classroom teachers and/or building principals implement the system to differing degrees. No "perceptions of the participants" are presented which could delineate perceived problems in adjusting, or accessing the system, perceived "locus of classroom control", or other features which may be viewed as salient and important to the users. These questions await further study.

Acknowledgements

I would like to acknowledge and express my appreciation to the following individuals for their invaluable help in the planning, implementation, analysis, and interpretation of this study. Without this cooperative effort, this study could not have been accomplished. Dr. Thomas Payzant, Superintendent of OKC Public Schools, Dr. Jesse Lindley, Asst. Superintendent Educational Services, Dr. Alice Houston, Director of Curriculum Services, Ms. Billie McElroy, PRI/RS Data Consultant. Thanks also to Ms. Pat Watson, Director of Planning, Research, and Evaluation, for guidance and criticism, Dr. John Crawford, Research Associate, and Ms. Maridyth McBee, Research Associate for their help in the analysis and interpretation of the data, and Mr. Frank Raia, Research Assistant, for his help in compiling the data into an analyzable format. I would also like to acknowledge Ms. Kay Schmidt, Graphic Services, for the elegant and professional graphs and figures, and finally, special appreciation to Ms. Jocelyn Ellis for preparation of this manuscript.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The Board of Education of the Oklahoma City Public Schools has established twelve goals and administrative objectives for 1981-82. They are based on the assumptions that (1) all children can learn, (2) all children can be taught, (3) every child should show achievement gains from September to May, and (4) there are specific, known factors that make schools more effective.

Among these Board goals and objectives, the following specifically address the instructional approach and philosophy of the Oklahoma City Public Schools:

> Goal 1 - The board will expect the staff to assess the needs of each child so that appropriate educational programs can be provided to meet the different needs of the slow, average and gifted learner.
>
> Goal 2 - The board will expect the staff to assist each student according to his or her ability to make progress each year in learning and applying skills in reading, writing and computation.
>
> Goal 4 - The board will continue to support the development, refinement, and implementation of a curriculum which fosters mastery learning of the basic skills and accommodates a criterion-referenced testing program that enables teachers and administrators to manage instruction and assess student progress.

As is obvious, individualization of instruction, and mastery of basic skills is the applicable philosophy, and the use of a valid and reliable criterion-referenced testing program is the curriculum orientation, management device, and feedback mechanism designated for this approach.

Congruent with these directions, in 1980 and 1981 the Educational Services staff evaluated both the available commercially published management

systems and the possibility of developing a custom-designed criterion-referenced testing system unique to Oklahoma City. After consideration of logistics, time referents, reliability and validity constraints, as well as expense, the decision was made to purchase the Prescriptive Reading Inventory Reading Systems (PRI/RS) published by CTB/McGraw Hill, Monterey, California.

Staff analysis indicated that the PRI/RS System:

1. Had a high degree of coordination (97% match of the objectives) with the district's established reading/language arts objectives.

2. Had been successfully used in other large school districts.

3. As a management device, it would help facilitate the coordination of instruction between the regular classroom teacher, Special Education and Title I teachers, and also facilitate the continuity of instruction between all teachers and schools within the district.

4. Had substantiated the validity and reliability of its criterion-referenced testing development procedures.

5. Had criterion-referenced mastery tests which were readily available and compatible with a computer management system (TRACER) already in use in the Oklahoma City Public Schools.

6. Had instruction, and support materials which were well planned, organized, skill defined, and which provided illustrated procedures for reading skill instruction and management.

Therefore, the PRI/RS was implemented during the 1980-81 school year in Grades K-5. All K-5 students in the district (except those not in attendance) were administered the PRI/RS diagnostic test, which provided an index of reading objectives they had already mastered (for their developmental level) as well as an index of those they had not mastered.

PRI LEVELS	GRADE EQUIVALENCY	INSTRUCTIONAL OBJECTIVES PER LEVEL
A	0.0 - 0.9	13
B	1.0 - 2.5	39
C	2.6 - 3.5	43
D	3.6 - 6.5	42
E	6.6 - 9.0	34

The PRI/RS management procedures and components are described as follows:

Part 1: Diagnosis

The diagnostic tests are scored, and the results reproduced in three informational report forms; the Individual Diagnostic Map, The Objective Mastery Report, and the Class Grouping Report. The Individual Diagnostic Map provides a record of each student's needs, listing the reading objectives, the level tested, and the student's performance on each objective. The Objective Mastery Report provides the teacher with a group listing of students who need instruction on specific objectives. The Class Grouping Report provides a class summary of each objective, so that teachers can immediately see areas of greatest need (or best achievement) of their students.

Part 2: Prescription

The Teacher Resource Kits contain instructional support materials which provide skill information, file cards, lesson plans, work sheet activities, tutorial activities, and mastery

tests for each of the system's 171 reading objectives.

Part 3: Teaching

The Teacher Resource Correlations provide a keyed reference source to reading, language and spelling adoptions. This allows for continuity of instruction in reading to be provided while supporting multiple adoptions.

Part 4: Monitoring

Mastery tests are used to assess mastery of objectives after instruction. There are 171 mastery tests, one for each instructional objective. The mastery tests are computer scored, thus providing immediate feedback regarding mastery or non-mastery to the teacher and student. With this information, the teacher can devise corrective instructional procedures to remediate the student's skill deficiency.

TRACER: Tracer is the software package utilized for scoring tests, recording and storing data, and generating records of progress and reports. Individual student acquisition data is stored by student number, and follows the student record through moves (or school transitions) to his/her new teacher.

The classroom teacher receives four major reports from TRACER; (1) the Student Mastery report, which indicates mastery, the need for review, or non-mastery of an objective, (2) Status II, which is a grouping report that assists in planning and grouping students who did not show mastery and need to be retaught, (3) Status III, which is a graphic representation of the progress of a class and (4) the Student Record Display, which lists all of

the reading objectives mastered by level, by an individual student.

Purpose of this Study

The purpose of this analysis is multifaceted. The questions to be addressed include: (1) "To what extent does degree of utilization affect comparable groups of students on standardized achievement measures?" (2) "Does degree of utilization differentially influence students' achievement based on their initial ability levels?" and (3) "To what extent does differential utilization of PRI/RS and TRACER affect reading acquistion measures such as number of objectives mastered on the diagnostic test, and retention of skills mastered (from both diagnostic to diagnostic test administration, and mastery test to diagnostic test administration).

Method

Impact of degree of utilization on Standardized Achievement scores:

Establishment of populations. Initial analysis of potential comparative student populations for the elementary grades (K-5) affected by implementation of PRI/RS in 1980-81, produced some immediate constraints. Analysis of kindergarden was not possible because 1980-81 was these students' first year in the OCPS system. No index of inital achievement abilities (pre CAT score) was available. Similarly, analysis of 1st grade was not possible because the Spring 1981 CAT test was the first year the instrument was administered to kindergarten children. Therefore, this group of individuals also had no measured index of skill acquisition prior to implementation of PRI/RS.

For grades 2-5, the initial population grouping consisted of differentiating high vs. low PRI/RS and TRACER utilization by classroom teachers for the months of January-April 1981. These months were selected because the degree of implementation district-wide was minimal prior to January, and the PRI/RS diagnostic test (which provides the matrix of mastery/non-mastery of reading objectives by level for the individual students) was administered in December. The end of April was the cut-off date due to the CAT administration early in May. May utilization was not represented in the analysis of effects on CAT scores because it would not impact or influence the CAT results.

The criteria utilized for differentiating groups based on high and low teacher utilization was: For low (1) mean TRACER utilization (and submitting student mastery tests) less than once per month (less than 4 times total), and (2) less than 50 student mastery test submitted (total) for the 4 month period. To be included in the high utilization group, it was necessary to have: (1) mean TRACER and mastery test submission of more than 3 times per month (more than 12 total), and (2) more than 100 total mastery tests submitted. These parameters allowed analysis of whether or not a sufficient number of students existed in both the high and low utilizations groups to permit stratified random sampling, with all ability levels represented in each group. Further, it facilitated the inclusion of a number of different teachers, and a number of different schools, which helped control for "teacher effects". In a small sample, with only a few teachers, this factor could be very influential in producing apparent differential treatment effects, independent of the actual treatment effects precipitated by PRI/RS and TRACER utilization.

Grade 3 had an insufficient number of high-utilization classroom teachers for comparison, and Grade 4 had an insufficient number of low-utilization individuals, therefore, these grades were eliminated from this analysis.

Table I describes the initial populations in Grade 2 and Grade 5 prior to the limitations imposed by the stratified random sampling procedures.

TABLE I

COMPARATIVE POPULATIONS BY UTILIZATION

GRADE	NUMBER OF TEACHERS	NUMBER OF DIF. SCHOOLS	TOTAL STUDENT POPULATION	$\bar{X}$ SUBMISSIONS PER MONTH	$\bar{X}$ TOTAL MASTERY TESTS SUBMITTED
2 High	8	5	173	6.8	>360
2 Low	8	6	167	0.78	<36
5 High	11	5	304	4.73	>306
5 Low	10	7	204	0.58	<26

A final caution needs to be inserted relative to the comparability and analysis of treatment effects by these groups. The differentiation of high and low utilization individuals does not imply that no effective teaching behaviors were occuring in the low utilization group. It is assumed that low utilization teachers were proceeding with classroom behaviors that were typical prior to implementation of PRI/RS, and were continuing to utilize management techniques and materials they were comfortable and familiar with, and that the students benefited from this treatment. Previous analysis revealed that CAT achievement gains in reading were demonstrated by students in both groups, and the results of this study do not rely on a comparison vs. a no-treatment control group. Rather, the intent is to interpret whether the high utilization of the regimen of PRI/RS criterion-referenced testing

program, and the individualized record of growth provided by TRACER, produces superior student achievement gain relative to "base line" data (i.e., the achievement of students taught under a traditional, non-criterion-referenced approach).

Establishment of the samples. For grades 2 and 5, the TRACER-generated student record displays of each student involved (both high and low groups) were reviewed to insure that the number of mastery tests submitted was consistent with the established parameters. A list of students in each group was established, and the Total Reading Composite raw score from the 1980 and 1981 California Achievement Test (CAT) were obtained for each individual. Any student not having both scores was eliminated from the population. For Grade 5, another constraint was imposed, because the special characteristics of this group provide further analysis of reading achievement independent of CAT score achievement analysis.

At the start of the 1981-82 school year, PRI/RS and DMI were initiated at the middle school level (Grades 6-8). Therefore, as before, the PRI/RS diagnostic tests were administered to these grades to determine degree of mastery/non-mastery of the objectives at these levels. This made the 6th grade students a special case in the district, as they were the only group to have taken a PRI/RS diagnostic test in December, 1980, as 5th graders, and again as 6th grade students in the fall of 1981. This provided the ability to assess the number of reading objectives mastered, independent of the CAT reading score analysis. Therefore, 5th grade students were eliminated from the population if they did not have both 1980 and 1981 CAT total reading scores, and results from the PRI/RS diagnostic test administered in December, 1980.

At this point, a new list was constructed, catagorizing individuals in each group by their demonstrated achievement level. The national percentile they achieved on the 1980 spring CAT Total Reading Composite score was utilized to group by increments of ten (i.e., 1st to 10th percentile, 11th thru 20th percentile, etc.). The results of their 1981 CAT performance were not included on this list. Within this stratified group, the sample was randomly selected without replacement with the following constraints: It must contain:

1. An equal number of students at each ability level (chance level thru 99th percentile)

and

2. Conform roughly to the district profile by sex and racial characteristics for that grade.

The sample eliminated students who scored at or below chance level +3 raw score points, because with chance level responding, it is unclear whether that student's score is an accurate indication of achievement. For Grade 2, chance level +3 on total reading is 21 (corresponding to the 9th percentile nationally). For Grade 5, chance level +3 on total reading is also 21 (which is the 14th percentile nationally on the spring norms.)

The final constraint involved individuals in the lower ability levels (below the 40th percentile nationally). To control for the effects of supplementary reading instruction via Title I, students scoring less than 40th percentile (national norms) were selected from non-Title I schools as much as possible. Although it was not possible to completely select non-Title I students and still maintain equal percentages across ability levels, the number of Title I students included was small (less than 5 per group), and was equated across groups (both Grade 2 and Grade 5).

Table II gives the breakdown of the final population grouping prior to stratified random sampling.

TABLE II

COMPARATIVE POPULATIONS AFTER CONSTRAINTS

GRADE	NUMBER OF TEACHERS	NUMBER OF SCHOOLS	REMAINING STUDENT POPULATIONS (N OF S's AFTER CONSTRAINTS)
2 High	8	5	97
2 Low	8	6	101
5 High	11	5	190
5 Low	10	7	111

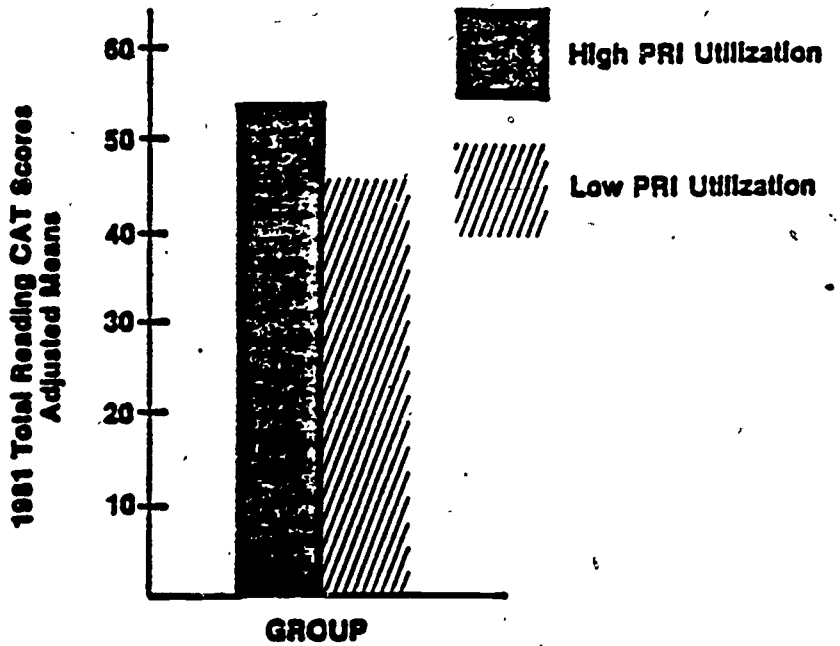
Results and Discussion

Grade 5

Effects of PRI/RS utilization on CAT Total Reading achievement scores:

Figure 1 gives the results of the Analysis of Covariance results between groups (high vs. low utilization groups raw scores) for Grade 5 students.

Figure 1
Grade 5-Main Effects



SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
COVARIATES:					
1980 ACHIEVE. SCORE	10348.87	1	10348.87	160.98	0.0
MAIN EFFECTS:					
GROUP	1186.10	1	1186.10	18.45	0.001*
EXPLAINED	11534.98	2	5767.49	89.71	0.0
RESIDUAL	6235.64	97	64.28		
TOTAL	17770.62	99	179.50		

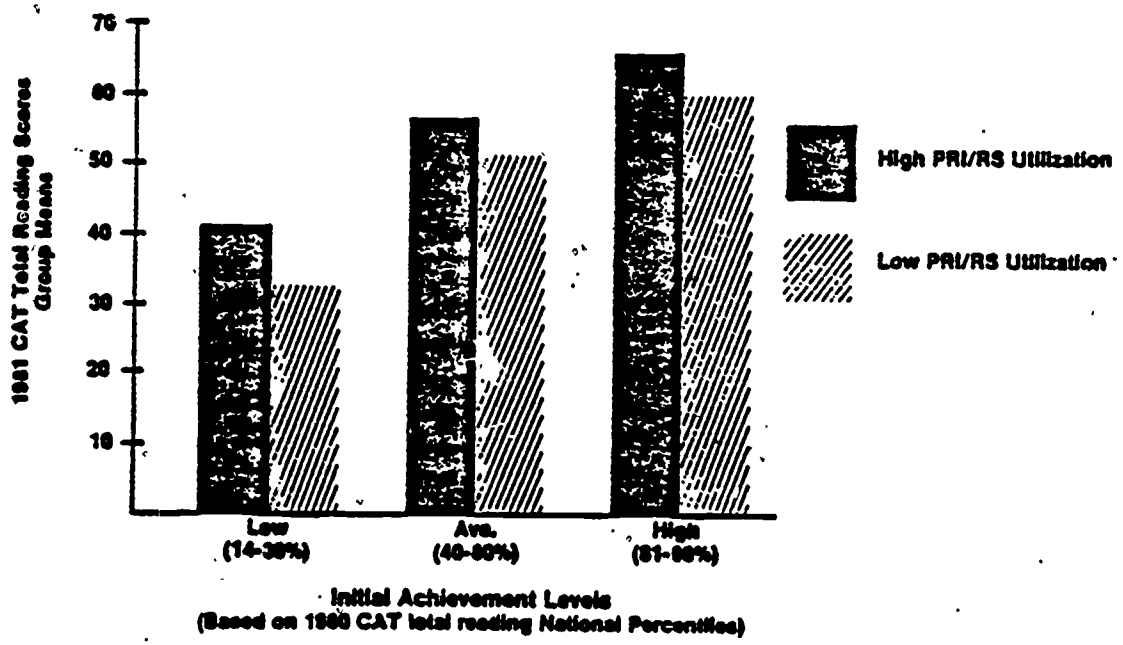
As Figure 1 shows, when the groups are equated for initial achievement differences (using 1980 Total reading CAT scores as the covariate), as a group, the students whose teachers exhibited a high degree of utilization of PRI/RS, TRACER, and mastery of objectives tests, scored significantly higher ($p < .001$) on the Total Reading Composite of the 1981 CAT.

As a supplementary analysis, it was of further interest to attempt to determine if this effect generalizes across ability levels, or if the between group differences are influenced (or caused by) differential achievement by portions of the groups (e.g., "Does the program benefit low-achievers vs. higher-achievers differentially?").

For this analysis, both groups were partitioned by their initial (before PRI/RS) achievement level. Students who scored between the 14th and the 39th percentile nationally on the 1980 CAT were designated as low-achievement individuals. Those scoring between the 40th and the 80th percentile nationally were designated as average-achieving individuals, and those between 81st to 99th percentile were designated as high-achieving individuals. Figure 2 gives the Analysis of Variance results of the 2X3 factorial design of groups

by achievement levels.

Figure 2
Grade 5-Simple Main Effects



SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
MAIN EFFECTS:	19851.69	3	6617.23	81.50	0.0
GROUP	799.52	1	799.52	9.84	0.002*
ABILITY LEVEL	19186.05	2	9593.02	118.15	0.001*
2-WAY INTERACTIONS:					
GROUP BY ABILITY LEVEL	40.88	2	20.44	0.25	0.77
EXPLAINED	19892.58	5	3978.51	49.00	0.0
RESIDUAL	7631.60	94	81.18		
TOTAL	27524.18	99	278.02		

The data in Figure 2 indicate that the utilization of PRI/RS effects are distributed equally across achievement levels. The significant ($p < .002$) difference between the groups is again reflected by the analysis, as well as the difference in achievement across ability levels. The parallel profiles indicate that PRI/RS utilization appears to influence CAT Total Reading achievement scores regardless of beginning achievement level, and the lack of a two-way interaction indicates that an unusual cell (for example: high utilization, low initial achievers) is not responsible for the between-group differences.

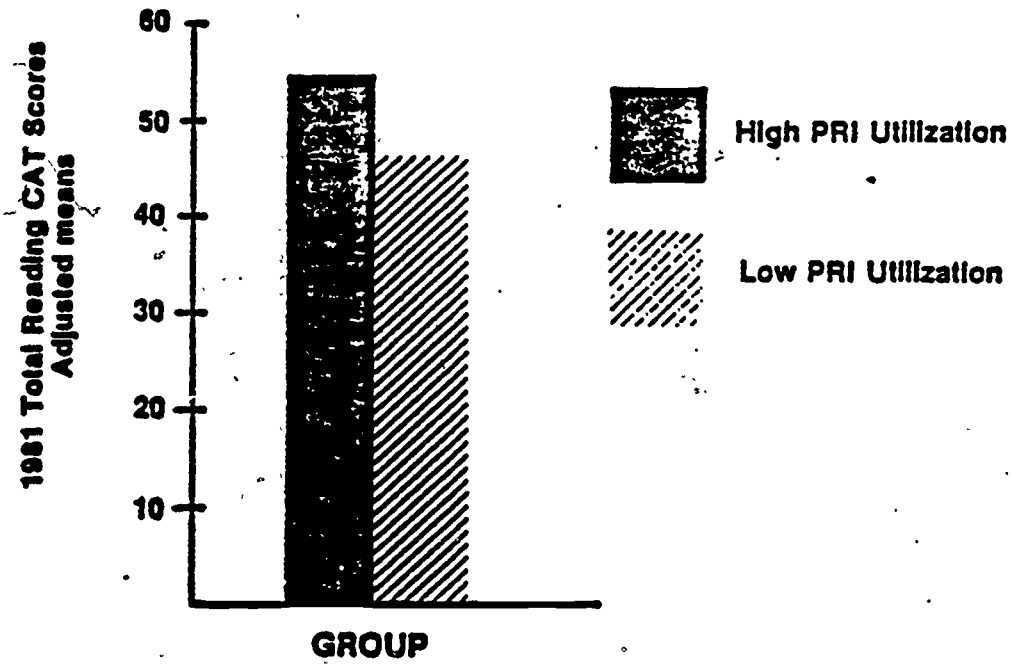
It is important to recall that these data do not imply that no learning occurred in the low-utilization group. What the data does show, is when the groups are equated on initial achievement measures (1980 Total Reading CAT National percentiles) and the group consists of an essentially equally distributed number of students across achievement levels, the utilization of PRI/RS, TRACER, Mastery tests, etc., enhances the achievement scores of the group of students. It does not appear to affect differentially high, average, or low ability individuals.

As an example, the high-utilization students who had scored between the 40th to 80th percentile nationally did not increase their mean scores relative to the low-utilization students who were also initially high achievers. They appear to maintain their relative position within-groups, but differential achievement is demonstrated between groups at all ability levels.

Grade 2

Figure 3 gives the results of the ANCOVA between groups for Grade 2 students.

Figure 3
Grade 2-Main Effects



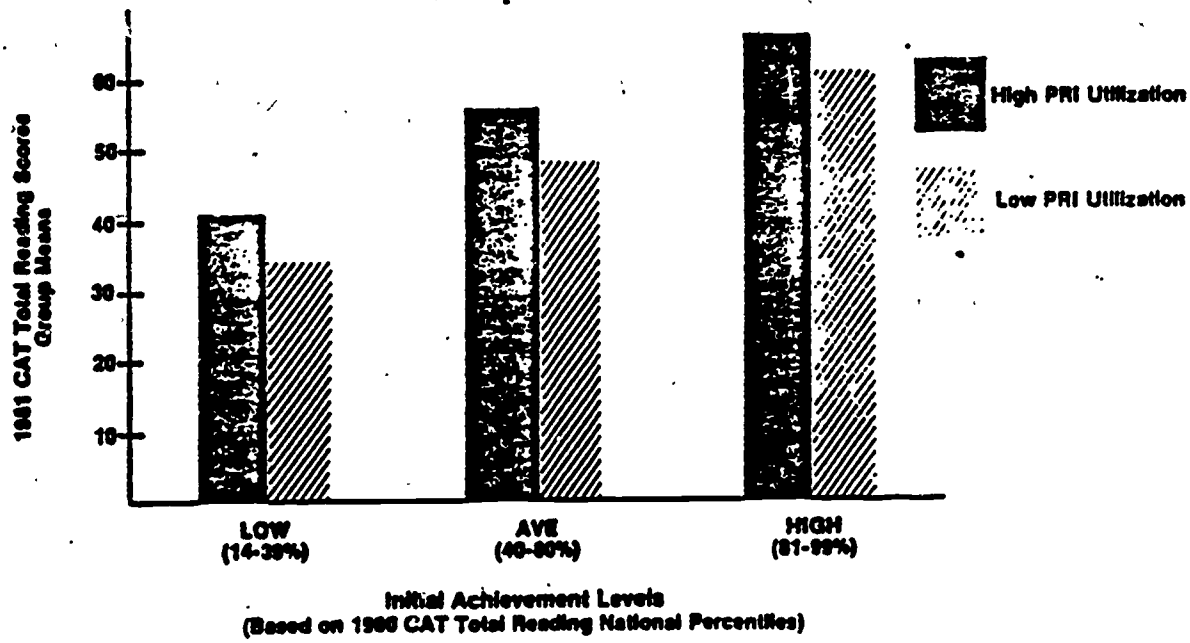
SQURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
COVARIATES: 1980 ACHIEVE. SCORE	10075.06	1	10075.06	94.77	0.0
MAIN EFFECTS: GROUP	1595.65	1	1595.65	15.01	0.001*
EXPLAINED	11670.71	2	5835.35	54.89	0.0
RESIDUAL	10311.91	97	106.30		
TOTAL	21982.63	99	222.04		

Grade 2 (cont.)

As Figure 3 shows, for Grade 2, when the pre-score covariate is utilized to reduce the effect of the initial achievement differences, as a group, the students experiencing a high degree of utilization of PRI/RS scored significantly higher ($p < .001$) on the Total Reading Composite of the 1981 CAT. This is consistent with the findings for Grade 5.

Figure 4 gives the breakdown by group and initial achievement levels for Grade 2.

Figure 4
Grade 2-Simple Main Effects



SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
MAIN EFFECTS:	11080.81	3	3693.60	31.90	0.0
GROUP	1162.06	1	1162.06	10.03	0.002*
ABILITY LEVEL	9440.56	2	4720.28	40.77	0.001*
2-WAY INTERACTIONS:					
GROUP BY ABILITY LEVEL	19.81	2	9.90	0.08	0.91
EXPLAINED	11100.62	5	2220.12	19.17	0.0
RESIDUAL	10882.00	94	115.76		
TOTAL	21982.63	99	222.04		

Again, consistent with the data on Grade 5, Figure 4 shows that the effects of utilization of PRI/RS are very nearly equally distributed across achievement levels for Grade 2. There is a significant difference between groups ($p < .002$) as well as the difference across achievement levels ($p < .001$). There is also no interaction between group and ability level, which indicates the effect is consistent across ability levels.

Summary and Conclusions

The data for Grades 2 and 5 appear to be consistent. When groups of students in these grades are equated for initial achievement levels (national percentiles on CAT) through a stratified random selection procedure so that all ability levels are represented, the statistical analysis indicates that those students participating in classes and schools where there is a higher degree of PRI/RS and TRACER utilization occuring, score significantly higher on the Total Reading Composite Score of the California Achievement Test.

However, in field research of this kind, as opposed to more rigorous laboratory research with stricter control of influential variables, the appearance of significant between-group differences does not automatically imply direct causal inferences. This is especially difficult to assess in the absent of randomized assignment to treatment and control groups, and a no-treatment control group for comparisons.

On the other hand, when the end result of the analysis is of a more applied nature (i.e., the ability of decision-makers to assess the direction of group effects, and the ability of a program to produce the "end-result" desirable behaviors such as improved reading achievement), the identification of combinations of influential variables which correlate highly with this desired result, are probably as "practically" useful as identifying single causal elements which influence theoretical learning parameters, retention, or retrieval characteristics of interest.

Assuming, therefore, that given the statistical assumptions, the random assignment within the groups, and the further precision of utilization of a covariate measure (to further statistically equate the groups prior to treatment effects), the results indicate that the utilization of PRI/RS (either as a powerful effect by itself, or in conjunction with other influential uncontrolled variables), produces statistically detectable differences in achievement.

The questions remaining are (1) What other uncontrolled variables could have produced these observed between group differences independent of PRI/RS utilization? (2) To what degree does the sample represent the OKC Public School population? and (3) How well do these presumed "learning effects" generalize to other measures of reading achievement gain?

Other Characteristics of the Sample

The first potential confounding variable which could influence the demonstrated between-group effects can be labeled "teacher effects". Although the sample was "spread-out" as much as possible to include as many teachers per grade (within the constraints of selection) which could be obtained, the selection by utilization still has inherent constraints. It is hypothesized that comparing groups of 10-11 teachers produces a relatively "normal" curve of teaching abilities for each group, and would negate the potentially powerful differences in effective teaching behaviors present when comparing student achievement gains across teachers. Nevertheless, it remains possible that along some dimension, as a group, the teacher variables may be functionally different. This does not assume the dimensions of competent/incompetent. Rather, all the teachers are assumed to be competent, and, given the same group of students, would be assumed to produce roughly equivalent learning. Instead, it is of interest to determine if the groups may have been different on measurable attributes such as length of experience, or specialist degrees. For example, it would be inappropriate to measure group outcomes on a measure such as PRI/RS utilization, if one of the groups varied dramatically in experience, or all had "Reading Specialist Certificates" or advance level degrees. Therefore, at least at a rudimentary level, it is important to compare these groups to determine if there are substantial and/or "obvious" differences.

Table III gives the breakdown by groups of experience and degrees of the selected teachers.

TABLE III

Teacher Attributes by Groups

Grade	N	Experience (Range)	Experience $\bar{X}$	Degree Obtained	Reading Specialist Certificate
2 Low Utiliz.	8	3-25 yrs.	13.5 yrs.	7-B.S. 1-M.S.	None
High Utiliz.	8	2-26 yrs.	13.6 yrs.	4-B.S. 4-M.S.	1 M.S.-Reading Specialist
5 Low Utiliz.	10	3-26 yrs.	9.8 yrs.	9-B.S. 1-M.S.	None
High Utiliz.	11	1-18 yrs.	9.8 yrs.	10-B.S. 1-M.S.	None

As Table III shows, the range of teacher experience in the Oklahoma City Schools are equivalent, with the mean number of years teaching experience nearly identical. For Grade 2, there may be an influence due to the high-utilization group having more advanced degrees, and one individual with an advanced degree and a reading specialist certificate. This could account for some of the between-group difference in achievement in grade 2. Although most past research has indicated that teachers' advanced degrees are not good predictors of subsequent student achievement, the Grade 2 results might reflect this difference, at least as a partial effect or as a combination effect with PRI/RS utilization.

For Grade 5 however, the groups appear to be functionally equivalent. There is no difference across groups in experience or obtained degrees, and none of the involved teachers have special certificates. This would indicate that the effect probably is not a particularly powerful effect, and if it is influential, it is only so in conjunction with PRI/RS utilization. Therefore, in the absence of more salient measures of teacher differences, it appears that the utilization of PRI/RS is a more powerful predictor of student achievement gains than the experience or degrees of the teachers.

Another question relating to the characteristics of the final randomly selected student groups, is: "Do they represent the district-wide race and sex distributions for that grade?" Again, analysis is necessary only to insure that grossly disproportionate representation does not qualify the generalizability to these grades district-wide. Table IV gives the breakdown by Grade.

TABLE IV

Student Sex and Racial Characteristics
of the Final Random Selection

Grade 2	High Utiliz	Low Utiliz	District
B	28%	24%	32.3%
W	64%	60%	56.4%
SP	4%	12%	4.8%
I	2%	4%	3.8%
O	2%	0%	2.5%
Sex			
M	48%	56%	51.9%
F	52%	44%	48.1%
Grade 5			
B	26%	32%	36.6%
W	56%	60%	52.8%
SP	2%	0%	4.3%
I	15%	6%	3.3%
O	4%	2%	2.6%
Sex			
M	50%	40%	51.9%
F	50%	60%	48.1%

As Table IV shows, there are slight discrepancies and variations from the district-wide averages (i.e., Blacks are slightly under represented while Whites and Indian students are slightly over represented). Yet any

perceived skew is not directional, and is certainly not grossly distored. So, for the purpose of this analysis, the samples can be considered to be representational. Again, as only as rudimentary measure, the discrepancies do not appear sufficient to produce the between-group achievement differences.

A final uncontrolled variable which may have impacted on the between-group differences is that of attendance. "Did the groups differ sufficiently in absenteeism to produce a substantive effect on the treatment conditions?" For this measure, the attendance data for January-May was obtained for each student in the sample. Table V gives the breakdown by grade and group.

TABLE V

Attendance

Grade	Number of Days Absent (Range) (Jan-May)	Median Number of Days Absent (Jan-May)
2 High Utiliz	0-21 days	5.0 days per person
Low Utiliz	0-53 days	6.5 days per person
5 High Utiliz	0-29 days	6.25 days per person
Low Utiliz	0-19 days	5.0 days per person

Analysis of attendance data reveals that for Grade 2 there may be a partial attendance effect, as the low utilization group had a higher median absentee rate per student than the high utilization group (6.5 days per person vs. 5.0 days per person). The median was utilized because one individual in the low group had 53 absences. Using the mean as a measure when an extreme score is present may distort the picture of the group-averages. However, in Grade 5, the achievement results do not reflect this influence, as the high utilization group had a higher

median number of days absent per student, yet still had significantly superior achievement scores.

Summary. In both Grade 2 and Grade 5, the degree of utilization of PRI/RS and TRACER appears to increase reading achievement significantly as measured by the California Achievement Test. In Grade 2, there may be additional influences based on teacher skills and differential attendance. This does not appear to be the case in Grade 5. These data appear consistent in their effect, and the sample appears to represent the district profile for these grades, permitting anticipation of improved standardized reading achievement scores district-wide as the system becomes more consistently and heavily utilized.

Effects of Utilization on Other Measures of Achievement

A final question remains, and relates to the generalizibility of the effects the utilization of PRI/RS has on reading achievement. Stated simply, the question is:

> "Does the demonstrated between-group difference (based on high and low utilization) on standardized achievement test scores, generalize to other measures of reading achievement, and reflect again the between-group differences based on degree of utilization?"

As mentioned earlier, these 5th grade students (in 1980-81) are a special case in the district. They are the only group of individuals who took the PRI/RS diagnostic test as 5th graders, and again this year (1981-82) as 6th graders. This, in effect, provided an independent pre-post measure on such variables as number of reading objectives mastered (from diagnostic to diagnostic test), the number of reading objectives retained (from mastery testing to diagnostic testing, and diagnostic to

diagnostic), and the proportion of retention and loss of the reading objectives (mastery demonstrated on pre but not on post). If the between-group difference is reflecting a true difference in reading achievement, the effects should replicate in this additional analysis, based on the same groups and degree of utilization of the system. As a further example, the high utilization group would be expected to have a higher number of mastery tests submitted, and a higher total number of objectives mastered by the very definition of the group. It is possible that the lower utilization teachers accomplished the same degree of reading objectives mastered, yet simply did not utilize the PRI/RS mastery tests or TRACER to quantify the gains. If there is a significant learning difference between groups, they should differ also on the new diagnostic test (assuming the students have taken "matched" Levels) on the number of objectives mastered. Further, one could speculate that although the numbers of mastery tests submitted would differ, the proportion of retention from diagnostic to diagnostic, and from mastery test to diagnostic should differ, if not significantly then at least in the same direction as previously demonstrated in the Achievement Scores analysis. To address these questions, an analysis of Reading Objectives mastered was appropriate.

Acquisition of Reading Objectives Analysis

Method

Groups. For this analysis, the same groups of 5th grade students were utilized. The PRI/RS diagnostic map (taken on 9-23-81 as 6th graders) for each student was obtained, and matched with the results on the diagnostic test from 12-5-80 (taken as 5th graders). Not every student in the

sample had taken the 9-23-81 diagnostic test, therefore, those who had not, were excluded from the analysis (four individuals in the high utilization group, and three individuals in the low utilization group were eliminated).

Measures. The nature of the PRI/RS diagnostic test, which quantifies mastery and non-mastery of specific reading objectives, makes possible a number of comparisons dealing not only with acquisition measures (achievement) but with retention measures as well. It is of interest to determine the effect of utilization on acquisition measures such as: (1) Total number of objectives mastered on 9-23-81 (hereafter designated as the "post-test"), and (2) Increase in the number of objectives mastered (pre-to-post).

For retention measures, it is possible to assess the effects of utilization on: (1) number of objectives lost (i.e. those specific reading objective that were shown mastered on the pre diagnostic test but were not subsequently shown as mastered on the post diagnostic test), (2) diagnostic test percentage of objectives retained (mastery demonstrated on both pre and posttesting), and (3) ratio or proportion of mastery tests submitted to diagnostic test demonstration of mastery. The use of ratios rather than raw numbers is necessary on the last measure due to the differential number of mastery tests submitted for the two groups. The very definition of the groups insures that more mastery tests were submitted per student in the high utilization group. The relevant question is "Did the students retain a higher ratio of the objectives, as measured by the post diagnostic test?"

Even given the fact that the groups were as closely equated for starting positions as possible, with all achievement levels represented,

there are a number of variables which could influence the between-group differences. These include: (1) Group (high vs. low teacher utilization, which based on the analysis of effects on CAT reading scores would be assumed to be a powerful influence on performance), (2) Initial starting abilities (in this case can be represented by two available measures, the 1980 CAT total reading composite, which was used to initially equate the groups, and the (3) Number of objectives the students had demonstrated mastery on the pre-diagnostic test.), (4) The number of days the student was absent (from Jan-May 1981) and (5) The level of the pre-diagnostic test (i.e. Level A-E).

Procedure. To compare the differential effects these five "predictors" (or independent variables, if they were manipulated in an ANOVA sense) have on the Dependent Variables (acquisition and retention measures) a Multiple Regression procedure was utilized.

To clarify multiple regression analysis, it is necessary to operationally define some terms. In a simple regression sense, the intent is to try and determine how effective a pre-measure (in this case an achievement score) is in "predicting" a subsequent post-measure of achievement (the dependent variable or measure). As an example, if the desire is to predict the national percentile in reading a student (or group of students) will achieve at the end of the school year, the strongest predictor is probably the national percentile they had achieved at the beginning of the year. Obviously, in the real world, there are other influences on learning, other than maturation. Attendance in school, what and how they are taught, and environmental factors all influence the degree of acquisition. In multiple regression, the

intent is to compare "multiple" predictors to the achievment (dependent) measure, to determine which predictor is the most powerful in predicting the actual gain. Further, it is of interest to determine if the other predictors add substantially to the predictive ability of the most influential predictor.

Results and Discussion

The first step in the analysis of multiple factors, and their potential influence on multiple dependent variables, is to compute a matrix of zero-order correlation coefficents. This gives a preliminary look at the degree of relatedness between the dependant variables and the predictors, as well as defining whether the relationships are positive (vary in the same direction) or inverse (where a high score on one measure is related to a low score on the other). It should be noted that the matrix of values represents all individuals in the study, without regard to high or low utilization of PRI/RS.

The matrix of the Pearson Correlations Coefficients is shown in Table VI.

The important facet of this analysis is primarily the relationship of dependent variable to dependent variable, and predictor to predictor. The relationship from predictors to dependent measures will be illuminated more clearly on the subsequent regression analysis. What this does show, is that there are several strong positive relationships among the dependent variables as well as some strong inverse relationships. The same is true of the predictor variables. As an example, the number of reading objectives the students mastered on the post diagnostic test are shown to be highly correl-

TABLE VI

PEARSON CORRELATION COEFFICIENTS-

		DEPENDENT VARIABLES					PREDICTORS				
		# of Obj. Mastered 1981	# of inc. Obj. Mast. 1980-1981	Diag. test percentage of Retain.	Mast. Test Ratio Retain.	# of Obj. Lost	CAT 1980 Tot. Read	Group (Util.)	# of Obj. Mastered 1980	# of Days Absent	Level of 1980 Diag. Test
DEPENDENT VARIABLES	# of Obj. Mastered 1981	———	———	———	———	———	———	———	———	———	———
	# of inc. Obj. Mastered 1980-1981	.61 P = .001	———	———	———	———	———	———	———	———	———
	Diag. Test percentage of Obj. Retain.	.82 P = .001	.55 P = .001	———	———	———	———	———	———	———	———
	Mast. Test Ratio Obj. Retain.	-.47 P = .001	-.44 P = .001	-.33 P = .013	———	———	———	———	———	———	———
	# of Object Lost	-.44 P = .001	-.62 P = .001	-.72 P = .001	.11 P = .23	———	———	———	———	———	———
PREDICTORS	CAT 1980 Total Reading	.36 P = .001	-.05 P = .35	.45 P = .001	-.25 P = .008	.14 P = .15	———	———	———	———	———
	Group	-.12 P = .13	-.29 P = .02	-.19 P = .09	.03 P = .37	.32 P = .01	.03 P = .38	———	———	———	———
	# of Obj. Mastered 1980	.28 P = .003	-.05 P = .36	.55 P = .001	-.28 P = .005	.03 P = .40	.50 P = .001	-.128 P = .10	———	———	———
	# of Days Absent	-.07 P = .24	-.26 P = .043	-.23 P = .05	-.05 P = .32	.13 P = .17	-.048 P = .32	-.217 P = .019	.08 P = .19	———	———
	Level of 1980 Diag. Test	.12 P = .12	-.15 P = .14	.12 P = .19	-.05 P = .29	-.06 P = .32	.59 P = .001	.235 P = .012	.16 P = .05	-.13 P = .11	———

ated with the number of reading objectives increased (between pre and post), and also highly correlated with the percentage of objectives retained (between pre-to-post). At the same time, the total number of objectives mastered on the posttest is inversely correlated with the number of objectives lost (between pre and posttesting). Put simply, this means the more reading objectives the student masters, the more he/she retains, and the less they lose between pre and posttesting. It is important to remember that these relationships do not illuminate group differences. Rather, they are representative of the 5th grade students in the sample as a group, without regard to differentiating high or low utilization of PRI/RS.

The predictor variables show the same mixed relationships. For example, the number of reading objectives mastered in 1980 (pre-diagnostic test) correlates highly with the CAT total reading pre score (the higher the CAT score, the more objectives they mastered in the pre-diagnostic test), while the CAT 1980 pre score shows no difference across groups (the groups are functionally equal on pre achievement measures). The CAT pre score also is highly positively correlated with the level of diagnostic test taken in 1980 (the higher the CAT score the higher the level of diagnostic test taken).

In order to get at the degree of influence of the predictor variables on the dependent measures, a step-wise multiple regression procedure was run for each of the dependent variables. This, in effect, tells the computer to analyze which of the predictors has the strongest influence on the dependent variable tested, and to insert that predicotr first in the regression equation. This will indicate how much of the variance in the sample is accounted for by that predictor. Then the next most influential predictor is inserted, and so on, until all are in the equation, and it is possible

to determine the total amount of variance accounted for, as well as the relative influence of each predictor. One caution needs to be addressed prior to analysis of the step-wise multiple regressions. In the case where large percentages of shared variances occur (such as the relationship between the number of objectives mastered on the pre diagnostic test and the CAT pre total reading scores), when one of them is inserted first into the model, the effect of the second one will be underestimated (because the variance which was shared is attributed to the first entry). Only the variance unique to the second measure will be reflected, and therefore its overall influence will be underestimated.

The summary of the analysis of the effects of predictors on each of the dependent measures is shown in Table VII.

The data in Table VII reflect several noteworthy results. First of all, when looking at the analysis of the number of objectives mastered on the post diagnostic test, the most influential predictor is the number of objectives the students had mastered on the pre test. This is logical. The pre-measure probably should be the strongest predictor, based on the total number of objectives mastered. Interestingly, the second most powerful predictor is the "treatment" group code (High vs. Low utilization). The significant F-ratio indicates that utilization of PRI/RS is a substantial predictor of number of objectives mastered (at least $p < .05$). The third most powerful predictor is number of days absent. For acquisition measures such as this, a finding like this appeals intuitively, and supports the pertinent literature of the effects of attendance. The level of the pre-diagnostic test and the CAT pre reading score are not significant predictors of this measure. However, it should be recalled that CAT '80 scores, and objectives mastered in '80

TABLE VII

STEP-WISE MULTIPLE REGRESSION SUMMARY

DEPENDENT VARIABLE: Number of objectives mastered on 9-23-81

PREDICTORS (in order of entry)	R^2	R^2 Change	F - Ratio
1. Objectives Mastered 1980 (pre-ability)	.515	.515	21.66 **
2. Group (High vs. Low utiliz)	.563	.048	4.80 *
3. Number of days absent (student)	.603	.040	4.67 *
4. Level of PRI Diag. Test (1980)	.610	.007	2.44 (NS)
5. CAT 1980 Total Reading Composite	.627	.017	1.77 (NS)

DEPENDENT VARIABLE: Number of objectives increased (12-5-80/9-23-81)

PREDICTORS (in order of entry)	R^2	R^2 Change	F - Ratio
1. Group (High vs. Low utiliz)	.102	.102	4.70 *
2. Days Absent (student)	.178	.076	4.17 *
3. Level of PRI Diag. Test (1980)	.192	.014	1.68 (NS)
4. CAT 1980 Total Reading Composite	.204	.012	0.90 (NS)
5. Objective Mastered 1980 (pre)	.211	.007	0.36 (NS)

DEPENDENT VARIABLE: Number of objectives Lost (12-5-80/9-23-81)

PREDICTORS (in order of entry)	R^2	R^2 Change	F - Ratio
1. Group (High vs. Low utiliz)	.132	.132	7.84 **
2. CAT 1980 Total Reading Composite	.192	.060	3.09 (NS)
3. Days Absent (student)	.226	.033	1.66 (NS)
4. Objectives mastered 1980 (pre)	.232	.006	0.79 (NS)
5. Level of PRI Diag. Test (1980)	.243	.011	9.58 (NS)

DEPENDENT VARIABLE: Diag. Test Percent of objectives retained (12-5-80/9-23-81)

PREDICTORS (in order of entry)	R^2	R^2 Change	F - Ratio
1. Objectives Mastered 1980 (pre)	.261	.261	5.96 *
2. Group (High vs. Low utiliz)	.323	.062	5.15 *
3. CAT 1980 Total Reading Composite	.377	.054	1.66 (NS)
4. Days Absent (student)	.418	.041	2.85 (NS)
5. Level of PRI Diag. Test (1980)	.418	.000	0.02 (NS)

DEPENDENT VARIABLE: Ratio of Mastery Tests Submitted to diag. test retention

PREDICTORS (in order of entry)	R^2	R^2 Change	F - Ratio
1. Objectives Mastered (1980 pre)	.175	.175	3.99 (NS)
2. Level of PRI Diag. Test (1980)	.191	.016	0.21 (NS)
3. Group	.195	.004	0.20 (NS)
4. CAT 1980 Total Reading Composite	.196	.001	0.04 (NS)
5. Days Absent (student)	.196	.000	0.02 (NS)

** ($p < .01$)
* ($p < .05$)

are highly correlated, so the effect of CAT '80 scores is underestimated due to the large degree of overlapping variance between those two variables. Were the experimenter to "force" the CAT '80 effects in first in the equation, this shared variance would be attributed to CAT '80 rather than objectives mastered '80, and show up as a more substantial predictor. For the purpose of this analysis it is sufficient to note that the pretest measure, the high vs. low utilization code, and attendance data are the significant predictors related to the number of objectives mastered on the post-diagnostic test. The R^2 refers to the total percent of variance accounted for in the sample, and the R^2 change is the contribution of each predictor to that total. To use the same example, 51.5% of the 62.7% total variance accounted for (in the total number of objectives mastered) is attributable to the pretest variance. Of the remaining 11.2%, 4.8% (or almost one half of the remaining variance) is attributable to the distinction between high and low utilization (or treatment differences), and 4% due to differential attendance between the students.

For the number of objectives increased (in mastery) between pre and post diagnostic test, the number of pre objectives mastered is not a substantial predictor. Again, intuitively this is straightforward. The number of objectives increased concerns only the number of total objectives mastered, minus the number duplicated on both tests, and minus the number lost (between pre-post). For this measure, the group and attendance figures are the only substantial predictors.

For the number of objectives lost (pre-post) the only significant predictor is the Group (degree of utilization).

For the precent of objectives retained between pre to post diagnostic tests, the most powerful predictor is again the number of objectives mastered

on the pre test. Again, the only other significant predictor is the Group (degree of utilizaticn).

As a final measure, the ratio of mastery tests submitted to retention on the diagnostic test, none of the predictors show strong relationships. The reasons for this are not clear, and cannot be determined with these data. It could be that the actual number of mastery tests submitted by the low-utilization group was so small, their ratio is spuriously high in comparison with the ratio of retention of the high utilization group. This cannot be determined empirically.

As a whole, it is striking that the "degree of utilization" predictor (the Group code) is the only consistent predictor to show up as being influential on all dependent variables (except the mastery test measure). And, as expected, the attendance factors and pre-ability measures are influential on the various dependent variables as well.

The Effects of Utilization of PRI/RS on the Dependent Variables

Because the degree of utilization has been shown to be an influential predictor on the various dependent measures, the next appropriate analysis is to attempt to determine the between-group effects high or low utilization of PRI/RS has on these variables.

Initially, since it has already been determined that attendance also influences these measures, and the previous CAT scores analysis revealed that the groups differed in attendance, it is necessary to determine if the groups differed significantly on attendance, so the effects of this factor can be reduced on the subsequent analysis.

For this analysis, a one way analysis of variance between-groups was utilized to test this assumption. Table VIII gives the breakdown between groups.

TABLE VIII

BETWEEN-GROUP ANOVA ON ATTENDANCE

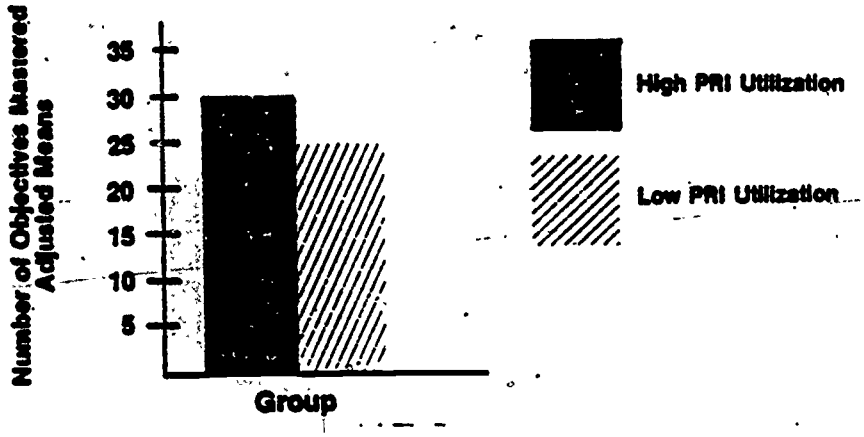
Group	N	Sum	Mean	SD	Sum of SQ.	Mean Sq.	df.	F	Signif. of F
High utiliz	46	396.0	8.6	6.76	155.6	155.6	1	4.48	.037*
Low utiliz	47	283.0	6.0	4.96					

As Table VIII shows, the 5th grade groups do differ on attendance, with the high-utilization group showing significantly ($p < .037$) higher absenteeism than the low utilization group. Therefore, whenever attendance was shown to be a strong predictor on the dependent variable, the days absent was utilized as a covariate to help reduce the effect of having differences between the groups on attendance.

The Effects of Utilization of PRI/RS on the Number of Objectives Mastered:

As was indicated earlier, the number of objectives mastered on the pre-diagnostic test, the group, and the number of days absent were all strong predictors of the total number of objectives mastered. Therefore, when analyzing the between-group differences, it was necessary to use as covariates: (1) the number of objectives mastered on the pre-test, and (2) the number of days absent. Figure 5 gives the results of the ANCOVA utilizing number of objectives mastered 1980 and number of days absent as covariates.

Figure 5
Number of Objectives Mastered
Post Diagnostic Test



SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
COVARIATES:					
NUMBER OF OBJECTIVES MAST. 1980	1979.85	1	1979.85	56.13	0.0
NUMBER OF DAYS ABSENT	126.81	1	126.81	4.23	0.045
MAIN EFFECTS:					
GROUP	132.51	1	132.51	4.42	0.04*
EXPLAINED	2236.87	3	745.62	24.90	0.000
RESIDUAL	1347.11	45	29.93		
TOTAL	3583.99	48	74.66		

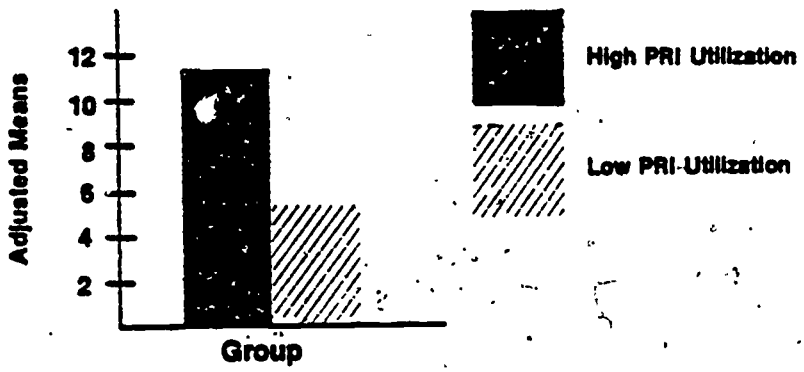
As Figure 5 shows, when number of objectives mastered on the pre test and attendance are used as covariates, the high PRI/RS-utilization group achieves significantly ($p < .04$) better as a group than the low PRI/RS utilization group. This is consistent with the analysis of the effects of utilization on CAT Total Reading Scores.

The Effects of Utilization of PRI/RS on the Number of Reading Objectives Increased (diagnostic-diagnostic test):

Since both group and attendance were shown to be strong predictors on the number of objectives increased in mastery between pre and post diagnostic

tests, ANCOVA was again utilized to determine the amount and direction of the between-group differences on this measure, with number of days absent as a covariate. Figure 6 shows the ANCOVA Results.

Figure 6
Number of Reading Objectives Increased
(Diagnostic—Diagnostic Test)



SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
COVARIATES:					
NUMBER OF DAYS ABSENT	90.84	1	90.34	3.78	0.058
MAIN EFFECTS:					
GROUP	128.31	1	128.31	5.34	0.025
EXPLAINED	219.15	2	109.57	4.56	0.015
RESIDUAL	1127.71	47	23.99		
TOTAL	1346.87	49	27.48		

As Figure 6 shows, the high-utilization group shows a significant ($p < .025$) difference in number of reading objectives increased (mastered) between pre and post diagnostic testing. This acquisition measure varies slightly (conceptually) from the total number of objectives mastered (see Figure 5). The increase in number of objectives mastered, deducts the number of objectives duplicated (mastered on both tests) and the number lost (mastered on the pre but not the post). Again, on this measure of achievement, the between-group differences are both consistent and in the same direction

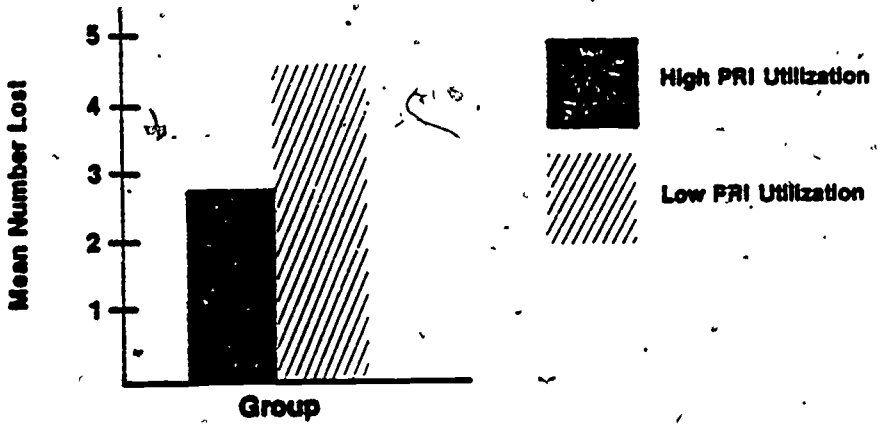
as the previous analyses.

Effects of Utilization on Number of Objectives Lost (pre-post):

For the analysis of the number of objectives lost between pre to post diagnostic tests, the only strong predictor was the group code. It is interesting that on this "retention" measure, unlike the previous acquisition measures, attendance is not a powerful predictor. It could be that "cognitive abilities" in conjunction with a "practice effect" (degree of utilization) are more important in retention or recall, while attendance has obvious implications for acquisition.

However, as group was the only powerful predictor, a one-way ANOVA was utilized to analyze the between-group differences on number of objectives lost. (See Figure 7).

Figure 7
Number of Reading Objectives Lost
(Diagnostic—Diagnostic Test)



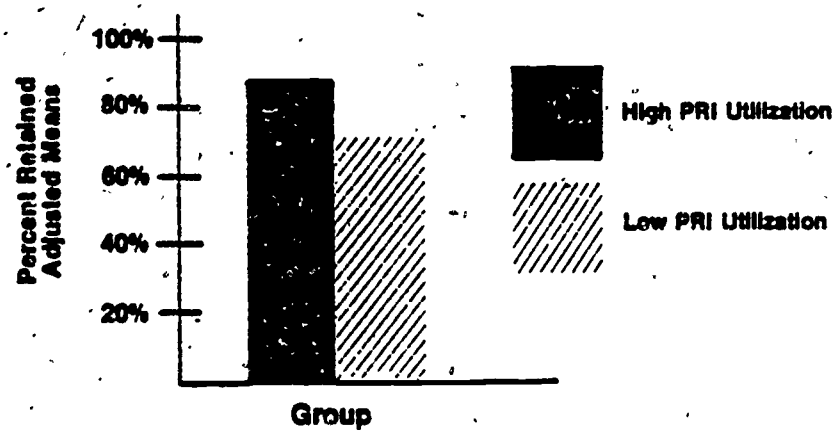
SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
BETWEEN GROUPS	36.07	1	36.07	5.79	.02*
WITH GROUPS	299.05	48	6.2		

As Figure 7 shows, the high utilization group had significantly ($p < .02$) less loss of objectives between pre to post diagnostic tests. This implies that degree of utilization affects not only acquisition results in reading achievement, but retention measures as well.

Effects of utilization on Percentage of Objectives Retained:

An equally salient measure of retention is the number of objectives duplicated in mastery performance between the pre-to-post diagnostic tests. For this analysis, the number of objectiver mastered on the pre-diagnostic test was shown to be a strong predictor, therefore it was utilized as a covariate in the ANCOVA. Figure 8 gives the result of this comparison.

Figure 8
Percentage of Objectives Retained
(Diagnostic—Diagnostic Tests)



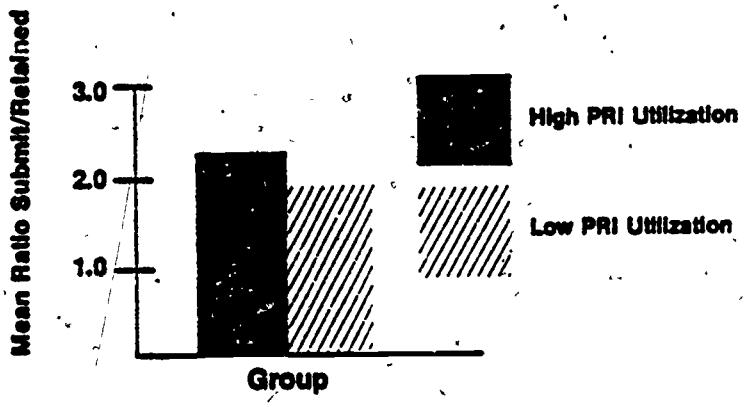
SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
COVARIATES:					
NUMBER OF OBJECT MASTERED 1980	0.396	1	0.39	21.86	0.0
MAIN EFFECTS:					
GROUP	0.068	1	0.068	3.78	0.058
EXPLAINED	0.465	2	0.23	12.82	0.0
RESIDUAL	0.852	47	0.018		
TOTAL	1.316	49	0.027		

Figure 8 shows that when number of objectives mastered on the pre-diagnostic test is the covariate, the high-utilization group retains a higher percentage of objectives than the low-utilization group. Although the F-ratio ($p < .058$) is not quite statistically significant at the .05 level, it is certainly "marginally" significant. Standing alone, this result would not be a particularly compelling argument of between-group differences. However, taken in conjunction with all the previous acquisition and retention analyses, it should be noted as being in the same direction.

Degreee of Utilization on Retention of Mastery Tests

The final analysis attends to the measure of the ratio of mastery test submissions, to the number of objectives subsequently shown to be mastered on the post-diagnostic test. As mentioned earlier, by the very definition of the groups, the number of mastery tests submitted are expected to be different. The ratio was utilized as a function of this difference. No predictor variable was shown to be particularly powerful on this dependent variable, so a simple one-way ANOVA between-groups was utilized. (See Figure 9).

Figure 9
Mastery Tests
Ratio of Submissions to Retention on Diagnostic Test



SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
BETWEEN GROUPS	0.75	1	0.75	0.32	0.57(NS)
WITHIN GROUPS	100.45	43	2.33		
TOTAL	101.20	44	2.30		

As Figure 9 demonstrates, the lack of a strong predictor is evident in the between-group differences. There is no difference statistically between the groups. Whether this is a function of the number of mastery tests being so low in the low utilization group, that it provides a spurious treatment effect (inflating the mastery ratio relative to the high-utilization group) or is functionally different along some dimension unique to the mastery test measure (rather than diagnostic to diagnostic tests measures utilized in previous analyses) cannot be determined with these data.

Summary

As with the data on the California Achievement Test results, the degree of utilization of PRI/RS appears to affect positively the acquisition and retention of Reading Objectives mastered. For acquisition measures, the

degree of utilization, the attendance of the student, and their starting abilities (as measured by the objectives already shown to be mastered on the pre-diagnostic test) were all strong predictors of achievement. For retention measures, it was shown that the degree of utilization of PRI/RS was the most consistent predictor of between-group differences, although pre-abilities were also a strong predictor for the number of objectives retained. Only the degree of utilization variable was consistent in being a strong predictor for both achievement and retention measures.

Conclusions

To summarize, and put into perspective the overall results of this study, it is necessary to remember that these designs can best be categorized as "quasi-experimental" designs. Without randomized assignment of subjects to treatment and control groups (with each subject having equal opportunity to be assigned to either), and without the added precision of a no-treatment control group for comparisons, the drawing of causal inferences about treatment effects is tenuous. However, given the a priori knowledge that the groups cannot be considered equivalent, every possible effort was made to equate (statistically and conceptually) the groups, so that perceived differences can be attributed to treatment effects. This involved attending to relevant variables which could not be controlled, and providing an index as to the effect these uncontrolled variables may have had on the demonstrated between-group differences.

It is further assumed that in studies such as this, where realistic knowledge of the effects of an instituted treatment has priority for decision-makers, over the ability to delineate single theoretical constructs (which

may be presumed to be the casual agents in the learning process) scientific analyses within the constraints of the data are appropriate.

Within the constraints of quasi-experimental design and analysis, it appears that the utilization of PRI/RS has a systematic and reliable effect on Reading Achievement. Those 2nd and 5th grade students whose teachers utilized PRI/RS and TRACER consistently and to a high degree, performed significantly better on the Total Reading Composite of the California Achievement Test. It also appears that the effect is distributed across ability levels. At all ability levels, the high utilization individuals demonstrated superior Reading Achievement results.

If this is a true learning effect, as it appears it may be, this has profound implications. One of the strongest implications is that unlike a hetrogeneous classroom situation where teachers might have the tendency to teach toward the "mean" ability individuals, the individualization and sequential record of growth for each child in the class (provided by PRI/RS and TRACER) allows the high achieving individuals acquisition at their own pace, as well as allowing low achieving students to maximize their learning.

The effects of utilization on Standardized Achievement Scores were consistent between Grade 2 and Grade 5. Further, the effects on Grade 5 achievement, appeared to replicate when other measures of achievement (acquisition) and retention were utilized. The degree of utilization appears to differentially affect the number of objectives mastered on the PRI/RS diagnostic test, the number of increase of reading objectives mastered, the number of objectives "lost" to mastery between pre-to-post diagnostic tests, and the percentage of retention of objectives between pre-to-post diagnostic tests.

In summary, even through the utilization of PRI/RS cannot be proven with this data to be the "only" casual agent producing these results, it certainly appears to be influential in the hypothesized direction, and is at least partially causal in conjunction with other uncontrolled or unidentified variables.

A final word of caution about the limited scope of this analysis. This research is an attempt to quantify and identify the effect utilization of PRI/RS may have on Reading Achievement, as the PRI/RS system becomes more functionally integrated and utilized system-wide in the Oklahoma City Public Schools. It is not, nor was it intended to be, an encompassing "Evaluation" of the PRI/RS system in the OKC Public Schools. No data are available to identify why some classroom teachers and/or building principals implement the system to differing degrees. No "perceptions of the participants" are presented which could deliniate problems in adjusting, accessing the system, perceived "locus of classroom control", or other features which may be perceived to be salient and important to the users. These questions await further study.