DOCUMENT RESUME

ED 131 092 24 TM 005 760

AUTHOR Leinhardt, Gaea
TITLE Program Evaluation: An Empirical Study of Individualized Instruction.
INSTITUTION Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
REPORT NO LRDC-1976-5
PUB DATE 76
NOTE 62p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS Academic Achievement; Achievement Tests; Aptitude Tests; Classroom Observation Techniques; Classroom Techniques; *Compensatory Education Programs; Data Analysis; Elementary Education; Federal Programs; Grade 2; *Individualized Instruction; Mathematics; Predictor Variables; *Program Development; Program Effectiveness; *Program Evaluation; Questionnaires; Reading; Standardized Tests; Student Records; Video Tape Recordings
IDENTIFIERS *Project Follow Through

ABSTRACT

The results of three years of evaluative investigation of the University of Pittsburgh Learning Research and Development Center's program of individualized education are described. The study was conducted in seven Follow Through sites and three Pittsburgh area schools. Standardized tests were used as input and outcome measures; questionnaires and videotapes were used to gather information about the classroom processes. The results indicate that over the three-year period the implementation of the program continued to move toward ideal goals. Extensive data reduction procedures were used and the rationale of usage discussed. Results also point to input as the primary explanation of student end-of-year performance, though classroom processes contribute a small but consistent amount. The data indicate greater ease in identifying negative factors in successful classroom processes than in determining the positive factors. (Author/MV)

LEARNING RESEARCH AND DEVELOPMENT CENTER

1976/5

PROGRAM EVALUATION: AN EMPIRICAL STUDY OF INDIVIDUALIZED INSTRUCTION

GAEA LEINHARDT

University of Pittsburgh

PROGRAM EVALUATION:
AN EMPIRICAL STUDY OF INDIVIDUALIZED INSTRUCTION

Gaea Leinhardt

Learning Research and Development Center

University of Pittsburgh

1976

The research reported herein was supported by the Learning Research and Development Center, supported in part as a research and development center by funds from the National Institute of Education (NIE), United States Department of Health, Education, and Welfare. The opinions expressed do not necessarily reflect the position or policy of NIE, and no official endorsement should be inferred. The author wishes to acknowledge the great assistance of Andrea Mar Seewald who helped with much of the data preparation, of Sundari Krishnamurthy who worked tirelessly at the computer, and of Chuck Teggatz who helped edit the manuscript.

Abstract

This study describes the results of three years of evaluative investigation of LRDC's program of individualized education. The study was conducted in seven Follow Through sites and three Pittsburgh area schools. Standardized tests were used as input and outcome measures; questionnaires and videotapes were used to gather information about the classroom processes. The results indicate that over the three-year period the implementation of the program continued to move toward ideal goals. Extensive data reduction procedures were used and the rationale of usage discussed. Results also point to input as the primary explanation of student end-of-year performance, though classroom processes contribute a small but consistent amount. The data indicate greater ease in identifying negative factors in successful classroom processes than in determining the positive factors.

4

PROGRAM EVALUATION: AN EMPIRICAL STUDY OF INDIVIDUALIZED INSTRUCTION

Gaea Leinhardt

Learning Research and Development Center
University of Pittsburgh

The area of program evaluation arises from the need to assess whether or not a specific policy or program has altered a trend or pattern of events in the desired way. The policy or program may be one designed to reduce or eliminate something such as crime or disease, or it may be designed to increase or generate something such as employment, educataional attainment, or the number of new minority businesses. In addressing the general problem of determining the effect of a particular policy or program, the specific problem of assessing the actual presence of a program in the field must be dealt with. To state it in a different way, the task is to establish a reasonable line of causality showing that a change occurred due to the presence of an implemented program. Assessing implementation is the focus of this paper.

Compensatory education is one example of a publicly funded program whose effectiveness requires empirical evaluation efforts. The idea of compensatory education is to compensate specific groups of children for an initial educational deficit by providing and funding special programs for them. Among the educational programs established in this way were Head Start and Follow Through. Head Start, as its name suggests, was designed to give economically disadvantaged children an early educational start at 3, 4, and 5 years of age. Follow Through was designed to protect and extend the presumed gains made in Head Start through the third grade of public education. It should be realized that these compensatory educational programs are not unitary. There is, for example, no single Follow Through program; there are at least 22 identifiable programs, and within the 22 there are probably several variations of sufficient magnitude to be considered as separate programs.

A separate federal program, the Cooperative Research Act of 1954, led to the establishment of over 40 research and development centers and educational laboratories. The research and development (R&D) centers were, for the most part, university-based organizations with specific "missions" to improve the quality of education in the United States. The educational laboratories, on the other hand, were established to refine and disseminate the findings and products of the R&D organizations to the educational community of the United States.

This paper presents empirical results from program evaluation undertaken at one of these R&D organizations, the Learning Research and Development Center (LRDC) at the University of Pittsburgh. The study is directed at evaluating LRDC's individualized instructional program in the context of the Follow Through system. Three concerns are addressed: (a) changes in levels of implementation of the major components of the program; (b) ways in which large amounts of intercorrelated data on classroom practices can be reduced and analyzed; and (c) the change or lack of change in the relation between classroom processes and student achievement.

Theoretical framework. It is important to view the research reported here in the light of its theoretical and practical roots, which lie in several rather separate areas. First, the motivation for the initial study grew out of the practical need for evaluation of LRDC's educational program in the Follow Through schools. Second, the techniques used to provide that evaluation came from two sources: observational literature and natural field-based research.

Most major models of program evaluation recognize the need for process information in order to interpret outcome results (Lindvall & Cox, 1970). Evaluative research, which aims to identify outcomes as well as explain the cause of outcomes, requires a reasonably detailed account of what the actual treatment was rather than what it was supposed to be (Charters & Jones, 1973). An additional requirement is that the treatment must be examined in an environment secure from extensive developer influence (Scriven, 1973).

In spite of the fact that the theoretical evaluation literature recognizes the need for process information, that literature itself is of little help in providing practical advice on its collection. This information has been gleaned instead from writers in the area of field-based research (e.g., Campbell, 1969), by examples from other evaluation researchers (e.g., Keeves, 1972; Stallings, Note 1), and from researchers in the area of classroom observation (e.g., Rosenshine, 1970).

Setting

The research was carried out during the 1973-74 academic year. It is a continuation and follow-up of research conducted during 1971-72 and 1972-73 (Leinhardt, 1974b). The evaluative effort has been focused at the second-grade level across several sites[1] rather than across grade levels at one or two sites on the assumption that variance is maximized within a grade level at different sites and minimized within a site.

---

[1] Site: A portion or all of a school district which has chosen an early childhood education program for use in at least one grade level in at least one school. In this paper, a site refers to at least one school using the LRDC model in a district. A local site is a school in the Pittsburgh area using the LRDC program without Follow Through funding.

Background. LRDC is one of 22 national Follow Through sponsors.[2] Each sponsor in the Follow Through system is responsible for implementing its own program at those sites which have chosen that particular model. LRDC implements its reading and mathematics programs in seven Follow Through sites and three local sites (for a brief program description, see Appendix A). This means that the programs are implemented to some degree in 61 second-grade classrooms. The LRDC Follow Through sites extend from North Dakota to Arkansas to West Virginia. The ten sites are located in urban, rural, and suburban environments; their student populations range from predominately white to predominately black, and one site is predominately native American Indian.

History. From 1969 to 1971, LRDC kept informal and often undocumented records of the implementation of its programs at the Follow Through sites. This effort was essentially an inventory of the presence or absence of vital materials or practices in the classrooms. Toward the end of 1970, it became clear that it would be useful to document program implementation on an expanded basis. Work was begun on defining basic dimensions of the program and constructing measures for each dimension (Leinhardt, 1972).

From 1971 to 1974, the evaluation project within LRDC[3] monitored the basic classroom processes that occurred in all second-grade classrooms in which some aspect of the program was implemented. The purpose of this monitoring effort was to assess the degree to which features

---

[2]Sponsor: An independent organization responsible for helping a site to implement an early childhood education program developed by the sponsor.

[3]LRDC is an organization of approximately 250 employees, 35 of whom are at the Ph. D level. The center is organized into 25 projects, each with one or more Ph. Ds and a support staff.

of LRDC's program were actually implemented in classrooms. The concern has changed, then, from a focus on an inventory to a focus on the degree to which the spirit of the program has been transplanted. This latter emphasis has itself gone through some changes and is moving toward the collection of information about classroom processes, independent of particular programs, that affect educational outcomes (Leinhardt, 1974b; Cooley & Leinhardt, 1975).

Procedure

During the 1973-74 school year, student performance and classroom process data were collected. These data were collected at seven Follow Through sites and three local schools, all of which were using LRDC curricula in math and reading. These data came from four sources: standardized tests, questionnaires, videotapes, and student records.

Aptitude and achievement data were obtained from three sets of standardized tests. The Lorge-Hagen-Thorndike Test of Cognitive Abilities (Thorndike, Hagen, & Lorge, 1968) was administered in the fall and served as an estimate of initial student ability. The Raven Coloured Progressive Matrices (1956) was administered in the spring and was one of three measures used to estimate student abilities at the end of the year. The Metropolitan Achievement Test (MAT) Form F or G (Durost, Bixler, Wrightston, Prescott, & Balow, 1971) was also administered in the spring and was used to estimate achievement in reading and mathematics (total subtest scores were used). These tests were then aggregated at the classroom level to estimate average class performance on the four standardized measures.

Data on classroom practices were obtained from responses to a questionnaire designed to yield information on a variety of topics. This questionnaire was completed by the educational specialist[4] at each site with the help, where appropriate, of the classroom teacher. (The questionnaire is reproduced in Appendix B.) The questionnaire was completed twice, once in the fall and once in the spring. The two responses were averaged, unless obvious errors were detected. This questionnaire generated data on 33 variables, of which 18 were retained for further analysis.

Data on the teaching practices in each classroom were obtained from videotapes. Each classroom was videotaped during reading and mathematics in the fall and spring. The tapes were made by the educational specialists and coded by the evaluation group at LRDC. Sixty minutes of videotape were recorded for each classroom. The analysis of these tapes yielded information on 30 variables, of which 20 were retained for further analysis.

Data on student performance within the curricula were obtained from the students' school records. Student initial and final location in both the reading and mathematics curricula were recorded and used in combination with the standardized test results.

The reliability of the instruments listed above has been estimated from two points of view: (a) the stability of the characteristics being measured over the year's period, and (b) the intercoder reliability. The stability of the variables measured by the questionnaire has been estimated at between .76 and .90 (Leinhardt, 1974b), and stability or generalizability of the variables measured by the videotapes has been estimated at .58. Intercoder reliability for the questionnaires and videotapes was estimated at .98 and .95, respectively.

---

[4] Educational Specialist: Individual(s) responsible at the site for implementing the sponsor's program(s).

Findings

In reporting the findings of this study, three questions are addressed. First, how has the implementation of a program of education changed over the course of three years? Second, how can the large mass of data generated by this investigation be reduced in a way which maintains the meaning but reduces the confusion of dealing with large numbers of variables and classrooms? Third, what is the impact of classroom processes on student achievement; how should these processes be considered, and how have they changed over the three-year period?

Implementation

In this section, the changes in classroom processes and the data which reflect them are examined. Before reviewing the details of the available data, some general comments are provided to serve as an overview.

From the 1971-72 school year to the 1973-74 school year, in accordance with the federal guidelines for Follow Through, the number of sponsor service days to the sites decreased. This is indicative of an initial policy decision which aimed at installing a program and then weaning the sites and sponsors from each other. Thus, the sites were expected to require and receive considerable aid initially in implementing the program, but they were expected to require and receive fewer and fewer supports as time passed. One would expect that if the program were difficult to implement or merely poorly implemented, there would be an increase over time of activities or behaviors that depart strongly from the sponsor's model. One would also expect that given a pattern of reduced support over time, teachers and specialists would begin to introduce their own innovations in answer to problems that arise. The question is whether or not those "answers" are in harmony with the spirit of the sponsor's model.

As the data on implementation are reviewed, two questions should be kept in mind: (a) Are key features of individualization maintained in the face of a reduction of support? (b) Does implementation increase over time in the face of increased experience with the program? Fourteen variables are presented in Table 1. These variables were selected from the many gathered each year when there were at least two years of data available on them. Table 1 shows the range of site means for the four sites on which we have been consistently gathering data. Each site has between four and eight second-grade classes.

Table 1

Mean Ranges for Classroom Process Information over Three Years with Original Four Sites

Variable	1971-1972	1972-1973	1973-1974
Teacher experience with LRDC program	.00 -- 1.83	.70 -- 2.20	1.50 -- 3.20
Class size	21.00 -- 27.40	20.60 -- 25.30	20.50 -- 25.80
Percentage of students attending	87.80 -- 96.60	91.00 -- 96.70	94.20 -- 95.80
Percentage of unique assignments	51.00 -- 57.00	77.00 -- 86.20	80.00 -- 85.50
Percentage of blocked assignments		47.30 -- 63.80	40.00 -- 73.50
Number of days since the last test	3.15 -- 9.06	3.40 -- 4.50	2.60 -- 4.20
Minutes per day in mathematics	44.16 -- 60.00	44.20 -- 60.00	60.20 -- 70.70
Minutes per day in reading	86.30 -- 120.00	76.00 -- 110.20	81.20 -- 98.30
Mathematics placement	6.35 -- 18.90	7.00 -- 20.90	9.80 -- 19.20
Reading placement	1.82 -- 4.87	3.40 -- 13.10	3.00 -- 9.20
Mathematics progress	6.26 -- 11.56	7.90 -- 9.50	9.00 -- 11.00
Reading progress	6.60 -- 6.86	5.00 -- 7.90	6.00 -- 7.70
Number of adults who travel	1.50 -- 2.30		1.50 -- 2.50
Percentage of negative statements	1.00 -- 6.00		.80 -- 4.00
Ratio of mean cognitive contacts to mean management contacts	.99		1.39

These variables are designed to provide information for the evaluative analysis. Since they are specific to this effort and are not generally known, they require some discussion as to meaning and utility. Rather than describing each variable in the table, a few of the more important ones will be reviewed. Attendance is often considered an indicator of student attitude toward school; high attendance indicates a more positive attitude and low attendance indicates negative attitudes. Regardless of whether or not that interpretation is warranted, attendance is important because it is unlikely that a student can learn much from school without being there. It is, therefore, encouraging to see that the lower end of the range of percentage of students attending has gone up from 87.8% to 94.2% attending.

Three variables--unique assignments, blocked assignments, and frequency of testing--relate to teacher assignment practices that are specified by the instructional model. One of these, unique assignments, will be discussed in detail. Individualized instruction requires that each student receive only instruction that he or she needs. It is assumed that if all students in a class simultaneously receive the same instruction, individual students will not be receiving exactly what they need. The percentage of unique assignments estimates how unique the entire group of class assignments are; the more unique, the higher the percentage. Clearly, it is not necessary for every child to receive a different assignment; but, an increase over time from approximately 54% to 83% unique is indicative of increasing individualization.

Four variables--math and reading placement and progress--relate to the initial starting position of children in the curriculum and their progress during the year. These variables are indicative of the impact of the instructional practices of not only the second grade, but also of the first grade. Both mathematics and reading placement have gone up over time. This is impressive because basic entering ability has remained stable, and higher placement levels thus indicate that more material is being taught in the first grade and retained in entering second grade. The students have made roughly similar progress over the three years in both subject areas, indicating that more material had been taught and, presumably, learned at the end of the second grade in 1974 than in 1972.

Several variables relate to [illegible] teaching practices in the classroom. One of these, the r[illegible] to management contacts, is indicative of the amount of emphasis placed by the teacher on instructional rather than managerial information. The larger the number, the more the teacher is interacting in an instructional way with students. This number has increased slightly over time. The LRDC instructional model seeks to minimize the dependence of children on teachers as managers, dispensers, and clerks. Its implementation should lead to a reduction in the number of managerial comments made, and the data support this contention.

Other variables in the table support a claim that the level of implementation of the LRDC program has remained stable during the three years in which support--in number of sponsor service days--was gradually being withdrawn. The number of adults in the room who "travel" (or circulate) during the working period has remained stable and higher than is indicated by the model. The percentage of negative comments made during an observed period of time has remained quite low and stable over the three years. The amount of time spent in teaching reading and mathematics has fluctuated, but is within developer guidelines, and tests are administered regularly in order to prescribe assignments effectively.

There is much, however, that is not revealed by these data. Over the three years that we have been observing teachers in classrooms using our programs, there has been a steady increase in the confidence with which the model has been implemented. The teachers and aides travel in the room, helping, tutoring, and quizzing students in order to diagnose their instructional needs. Specific program sequences are frequently altered on the basis of the accumulated wisdom of teachers and educational specialists. Different patterns of classroom organization are constantly being tried. Some teachers travel among the students as they do their prescribed work while the aide tests or engages in extended tutoring. In other situations, both the teacher and the aide travel among the students for a while to get them started on their separate t[illegible]d then engage in relatively brief, small group instruction. In stil[illegible] other cases, the teacher starts with small group instruction and follows that by joining the aide in traveling. In many cases, teachers have students help one another. Common to all of these patterns of interaction, however, is an attempt to implement the idea of individualized instruction. Over time, these four sites have increasingly captured the spirit of individualization. In those areas which are critical to the implementation of the instructional model, the sites have maintained or improved their practices. In those areas in which precise implementation is not as critical, the sites have expressed their own uniqueness.

Data Reduction

In the process of assessing the implementation of the instructional program, an enormous quantity of intercorrelated data is generated. Over the last three years, a haunting problem has been the reduction of the large number of process measures for purposes of a more complete analysis. There are two reasons why it is necessary to reduce the mass of classroom process data. First, there is the obvious problem of limited degrees of freedom for any analysis which requires that the number of variables be reduced by approximately a factor of 10. The second need for data reduction arises from the problems of trying to understand and interpret what all of the numbers mean. The trade-off is between large numbers of specific, concrete, easily understood bits of information and a total system of information. Although what is needed is to move toward a system of information, it often appears that detailed and specific information is of greater value for the practitioner. This problem has plagued most classroom researchers. It is common in almost any situation in which one attempts to relate detailed descriptive information to outcomes and in which one uses the classroom as the unit of analysis (e.g., see Stallings, Note 1).

One way to solve [illegible] n of degrees of freedom is to change the unit of analysis fron [illegible] ssroom to the student. It could easily be argued that in an individualized program the individual student is a more appropriate unit of analysis (e.g., see Brophy, Note 2). However, there are two reasons why it is inappropriate to consider the student as the unit in this case. First, the classrooms are self-contained, and although different students receive different "treatments," the students in a single class can hardly be said to vary independently of one another with respect to that treatment. Second, stable and reliable information on the individual student's sets of experiences is very costly to collect, requiring a minimum of two to two and one-half days of observations per child. Thus, using the classroom as the unit of analysis has permitted us to collect sufficient information to estimate the stability of procedures, the reliability of coding, and so on (Leinhardt, 1974a).

Data reduction: Possible approaches. Basically the problem of reducing data is the following: How can the number of variables be decreased without losing potentially valuable information? Clearly, the first step is to eliminate from consideration measures which cannot provide information for further analysis. This includes measures for which there is no difference among classrooms or for which there are obvious inaccuracies in the data reported. It should be remembered, however, that no variance in a measure does not mean that the measure has no information.

For example, if one finds that every child in every classroom has a mathematics textbook, that is important information. However, that particular variable in that particular sample will not be useful in predicting student achievement because there is no variance. After measures with no variance or with obvious errors are eliminated, the next step is to see if two or more of the measures are measuring the same thing. This information is obtained by examining correlation and partial correlation (with student entering abilities controlled) matrices. If two or more variables correlate highly (above .80) with each other and correlate to the same magnitude with the dependent [illegible]), it is possible that they are measuring similar phenomena. [illegible] can then be combined, or one of the set can be substituted for the others. These two procedures resulted in the elimination of 26 measures in the current data set.

One way of accomplishing the second step is to use some type of factor analysis of the process data. This technique has been successfully employed by Soar and Soar (1972) and others. My experience, however, has been that factor patterns tend to be illogical, that is, the variables that go together mathematically do not do so intellectually and do not replicate very well from year to year (Leinhardt, 1972). In [illegible] past, therefore, either the specific correlation or partial correlation of single variables with dependent measures has been reported (Leinhardt, 1974a), or the data have been combined using a different approach [illegible]einhardt, 1974b).

Another approach to combining variables is to use a model of classroom process and fit variables to the constructs of the model post hoc. The Cooley-Lohnes model (Cooley & Lohnes, in press), an expansion of the Carroll model (Carroll, 1963), has been especially useful for this purpose (see Leinhardt, 1974b; Stallings, Note 1; Cooley & Emrick, Note 3). The basic approach in using the Cooley-Lohnes model is to select variables which can logically be assigned to each of the model's six constructs and then to combine the variables within each construct. Combining the variables is accomplished either by factoring within the construct and summing using the factor scores or by transforming the raw data to standard scores and summing using equal weights. The latter approach was tried on the 1972-73 data set after the original data set had been reduced to six constructs, two of which related to student performance and four of which related to process. They were analyzed using commonality analysis (Mood, 1971). While this approach is extremely helpful in simplifying a large and complex data set, it is not completely satisfying if the data are collected with some other organizational scheme in mind and then are fit to the model post hoc. Any data analysis model is most usefully employed if the data are initially collected with the model in mind (see Cooley & Leinhardt, 1975).

<u>Data reduction: Current strategies.</u> Currently, the most useful approach seems to be to organize the data around some set of logically conceived clusters--such as the basic domains employed in developing the instrumentation--and then to examine each cluster separately for purposes of data reduction. Table 2 identifies the clusters, the specific variables in each cluster, and the procedure(s) which were employed to combine the measures. Initially, all of the measures were converted to standard scores (zero mean and unit variance). Then each cluster of variables was analyzed using principal component analysis with one or two factor solutions. These factors were then examined and used as guides for combining the data. If the factor analysis suggested that the variables equally related to a factor, we summed the scores without weighting them, making certain that "favorite" variables were retained. Equal weights were used for combining the following clusters: Input, Traveling Frequency I and II, and Outcomes. If the factor analysis suggested that the variables were differentially related to a factor, the weights of the variable on the factor were used to sum them. This was done for the Opportunity and Ratings of Instructional Behavior clusters.

Table 2

Map of Total Data Reduction

INPUT	Input (Summed Z scores, equal weighting)	Lorge Thorndike Cognitive Abilities Test Initial placement mathematics Initial placement reading
PROCESS	Process Set 1	
	Ratings of Instructional Behavior (Summed Z scores, weights suggested by factor patterns)	Active responses sought Modeling of response Focusing attention Contingent praise General praise Clear tutoring Class management Backward chains Refers to earlier information Solicits opinion of correctness Refers to earlier success
	Opportunity (Summed Z scores, weights suggested by factor patterns)	The number of adults traveling in mathematics The number of adults traveling in reading Availability of trained alternate Percentage of children present Time estimate Minutes a day spent using LRDC programs Minutes a day in large group activities
	Traveling Frequency I (Summed Z scores, equal weighting)	Reading cognitive statements Reading cognitive questions Reading tutorings Child initiated Proportion of cognitive statements to total statements
	Autonomy II (Summed Z scores, equal weighting)	Children decide what unit of math to study Children request special tutoring Children begin work as soon as they come in Peer tutoring occurs
	Assignments	Percentage of unique assignments
	Process Set 2	
	Autonomy I (Summed Z scores, equal weighting)	Children get own assignments Children get own prescription ticket Children decide when to take a test Children move back and forth between work and free activity
	Traveling Frequency II (Summed Z scores, equal weighting)	Percentage of negative statements during reading Percentage of negative statements during mathematics Number of errors made during reading Number of errors made during mathematics
	Testing	Number of days since the last test
	Process Set 3	
	Context	Teacher experience with the program
OUTCOME	Outcomes (Summed Z scores, equal weighting)	Ravens Progressive Matrices MAT reading total MAT math total End-of-Year reading placement End-of-Year mathematics placement

Following all of these steps succeeded in reducing the ratings of teacher behaviors to one variable and the frequencies of specific behaviors to two variables, to cite the most extreme cases. This improved the situation by leaving 10 clusters on which to predict one; however, with approximately 10% attrition in the orginal number of observations due to missing data, there were still too many variables for the available degrees of freedom.

A correlation was run on the 11 variables followed by a partial correlation with input partialled out to suggest further reductions. Tables 3 and 4 show the results of these analyses. After examining the correlations and partial correlations between the process variables and outcomes, the next step suggested by the data was to break the process variables into three groups:

Table 3

Intercorrelations Among Clusters for 1973-74 Data

				Traveling Frequency		Autonomy				Days Since	
	Input	Ratings	Opportunity	I	II	I	II	Teaching Experience	Percentage Unique	Last Test	Outcomes
Input	1.00										
Ratings	.01	1.00									
Opportunity	.10	.35	1.00								
Traveling frequency I	.23	.07	-.08	1.00							
Traveling frequency II	-.23	-.41	-.19	-.20	1.00						
Autonomy I	.13	-.14	.11	-.06	-.00	1.00					
Autonomy II	-.13	.35	.23	.12	-.14	.15	1.00				
Teaching experience	.61	-.08	.02	.00	-.03	.10	-.08	1.00			
Percentage unique	.09	.24	.08	-.02	-.32	.05	.08	.10	1.00		
Days since last test	-.14	.17	-.21	.13	-.06	.06	.09	.04	-.18	1.00	
Outcomes	.91	.12	.20	.27	-.32	-.01	-.01	.49	.16	-.29	1.00

Note: n = 55.

Table 4

Partial Correlations Among Clusters for 1973-74 Data

	Ratings	Opportunity	Traveling Frequency I	Traveling Frequency II	Autonomy I	Autonomy II	Teaching Experience	Percentage Unique	Days Since Last Test	Outcomes
Ratings	1.00									
Opportunity	.35	1.00								
Traveling frequency I	.07	-.10	1.00							
Traveling frequency II	-.42	-.17	-.16	1.00						
Autonomy I	-.14	.10	-.09	.03	1.00					
Autonomy II	.35	.25	.16	-.18	.17	1.00				
Teaching experience	-.11	-.05	-.18	.14	.03	-.00	1.00			
Percentage unique	.24	.07	-.04	-.31	-.07	.09	.06	1.00		
Days since last test	.17	-.20	.16	-.10	.08	.07	.15	-.17	1.00	
Outcomes	.28	.27	.16	-.28	-.30	.26	-.20	.18	-.40	1.00

Note: n = 55.

those which related positively to outcomes, those which related negatively to outcomes, and those which switched in their relationship to outcomes when input was partialled out. The autonomy variables, which are uncorrelated with outcomes, were assigned to one of the three groups based on the partial correlations. This organization of the data is also presented in Table 2; Process Set 1 consists of those five variables which relate positively to achievement, Process Set 2 consists of the three variables which relate negatively, and Process Set 3 consists of the single variable which changed after partialling from positive to negative. This stage of the data reduction can be considered a crude form of criterion scaling.

Process Set 1 became the sum of the following clusters: ratings of teachers, opportunity, traveling frequency I, autonomy II, and percentage of unique assignments. Process Set 2 became the sum of the following clusters: autonomy I, traveling frequency II, and the number of days since the last test. Process Set 3 is the single variable, the number of years of teaching experience with the program. The last variable is interesting because while teaching experience is correlated with outcome, it is more highly correlated with input, indicating that more experienced teachers are working in classrooms with more gifted students; but judging from the partial correlation results, they are not any more successful in teaching.

Data reduction: Future. In the future, it is hoped that the data will be collected in a way which will permit more logical and parsimonious reduction procedures. However, there will still be several steps involved.

1. The data should be checked to eliminate variables with little or no variability or measures with high numbers of obvious errors.

2. Preliminary correlations should be examined to see how the variables relate to each other.

3. Each variable should be examined to see if a tranformation is necessary in order to make the variable more additive with respect to the dependent variables. The exploratory data analytic techniques pioneered by Tukey (1970) and others are particularly relevant to these problems.

4. Logical clusters of the process variables should then be formed with the aid of factor analytic techniques when appropriate.

5. Preliminary correlations and partial correlations should be run on the reduced data set.

In summary, a particularly difficult problem in any research effort involving large numbers of variables and limited numbers of observations is the reduction of data. Over the past three years, we have evolved a system of exploratory data analysis to help in data reduction which keeps as much of the original information as possible, while reducing the absolute number of variables with which one is dealing. It has been our experience, however, that practitioners in the field are still most interested in the relationship between specific practices and student achievement. Therefore, it is important to have specific information easily retrievable and to display the data reduction process.

Analysis

This section considers how the impact of classroom processes on student achievement has changed over the three-year period, how this problem should be viewed, and the directions for future research. The basic question is: What is the influence, if any, of the educational process on student achievement? One of the more convincing methodologies used in recent studies in this area is commonality analysis (see, e.g., Mayeske, 1970; Mood, 1971; Cooley & Emrick, Note 3). Commonality analysis is an extension of multiple regression in which the total variance explained by the model is partitioned into variance which is unique to each predictor in the model and variance which is common to the group of predictors.

Results. Tables 5 and 6 present the results of the correlations and partial correlations of the reduced data set. From these tables, we can see that it is easier to determine what not to do in a classroom (Set 2) than what to do in a classroom (Set 1). We can also see that this procedure of reductions has substantially reduced the traditional problem of collinearity among the data.

Table 5

Correlations Among Sets for 1973-74 Data

	Input	Set 1	Set 2	Set 3	Outcomes
Input	1.00				
Set 1	.15	1.00			
Set 2	-.14	-.31	1.00		
Set 3	.61	-.04	.05	1.00	
Outcomes	.91	.29	-.33	.49	1.00

Note: n = 55.

Table 6

Partial Correlations Among Sets for 1973-74 Data

	Set 1	Set 2	Set 3	Outcomes
Set 1	1.00			
Set 2	-.30	1.00		
Set 3	-.17	.17	1.00	
Outcomes	.38	-.50	-.20	1.00

Note: n = 55.

If we analyze this reduced data set using commonality, the results indicate that none of the process data taken separately is significant in explaining achievement. However, the results of an analysis in which all of the process data are combined and considered as a unit are more promising. Table 7 shows that 88% of the variance in end-of-year achievement is accounted for by input and process information combined. Input alone explains a total of 83% of the variance, 47% of which is unique to input, while 36% is common with process. Process alone explains a total of 41% of the variance, 5% of which is unique to process.

Table 7

Commonality Analysis

	Input	Process
Total	.83	.41
Unique	.47	.05
Common	.36	.36

Note: R^2 = .88; n = 55.

In comparing this to previous years, it is most important to consider the zero order correlations between student input and outcome. These correlations have steadily increased since 1972 from .60 to .80 to .91 as a result of the increased precision with which entering and end-of-year abilities are assessed. The total R^2 has also increased from .68 to .72 to .88. The variance which can be uniquely attributed to classroom process has gone from a high of 14% in 1972 to 7% in 1973 and 5% in 1974. Undoubtedly, this has been adversely affected by the increasing size of the relationship between input and outcomes and by the limitation of using one educational program. This can also be explained by the fact that as classroom processes become more and more stable over time across grade levels, the power of entering student performance to explain end-of-year performance is likely to increase.

This analysis has several implications. First, initial student abilities consistently dominate the explanation of end-of-year student achievement. Second, classroom processes are contributing a small but consistent amount to the explanation of student achievement. Third, we are not yet ready to use these techniques to help identify the contribution of specific classroom processes which affect student achievement, not only because of the limitation of sample size, but also because of the examination of only one educational program. Currently, the best we can do is to point to practices which are consistent in their relationship to student achievement. Precision of assignment procedures, tutorial practices, and appropriate use of opportunities seem to improve student achievement, while negative and erroneous tutoring reduces student achievement.

Discussion

In this paper, I have argued that public policy evaluation requires that one include information about the nature of the program or the innovation rather than engaging in simple black box contrasts. In the process of measuring implementation, large amounts of intercorrelated variables are generated. In order to use the data in further analyses, they must be substantially reduced. Approaches to data reduction are, therefore, outlined. Finally, by using a form of regression analysis, outcomes are analyzed by taking into consideration not only initial input and programmatic descriptors, but also the interrelationship between the two. It is a fact of social innovations that the innovation is confounded with consumer self-selection, not only in education, but also in health care and other areas. Stated differently, social reforms are rarely randomly distributed; therefore, techniques which take into account and describe the nature of the confounding must be used for evaluative analyses.

This paper has used an empirical study of a particular educational innovation to demonstrate these points. In order to improve this particular line of research, several steps must be taken in the future. The data must be gathered according to a theory both to improve the interpretability of the results and to make the requisite data reduction more logical. The number of classrooms on which data are collected should be increased. And finally, the process information should be gathered from several programs instead of just one.

Reference Notes

1. Stallings, J. An implementation study of seven Follow Through models for education. Paper presented at the meeting of the American Educational Research Association, Chicago, April 1974.

2. Brophy, J. The student as the unit of analysis. Paper presented at the meeting of the American Educational Research Association, Washington, April 1975.

3. Cooley, W. W., & Emrick, J. A. A model of classroom differences which explains variation in classroom achievement. Paper presented at the meeting of the American Educational Research Association, Chicago, April 1974.

References

Campbell, D. Reforms as experiments. American Psychologist, 1969, 24, 409-429.

Carroll, J. B. A model of school learning. Teachers College Record, 1963, 64, 723-733.

Charters, W. W., Jr., & Jones, J. E. On the risk of appraising non-events in program evaluation. Educational Researcher, 1973, 2, 5-7.

Cooley, W. W., & Leinhardt, G. The application of a model for investigating classroom processes. Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1975. (Publication 1975/24)

Cooley, W. W., & Lohnes, P. R. Evaluative inquiry in education. New York: Irvington Publishers, in press.

Durost, W. N., Bixler, H. H., Wrightston, J. W., Prescott, G. A., & Balow, I. H. Metropolitan Achievement Tests (Primary II/Forms F & G). New York: Harcourt Brace Jovanovich, 1971.

Glaser, R. Evaluation of instruction and changing educational models. Los Angeles: Center for the Study of Evaluation of Instructional Programs, 1968. (CSEIP Occasional Report #13; also available as LRDC Reprint 46)

Keeves, J. Educational environment and student achievement. Melbourne: Australian Council for Educational Research, 1972.

Leinhardt, G. The boojum of evaluation implementation: Some measures. Unpublished doctoral dissertation, University of Pittsburgh, 1972.

Leinhardt, G. Observation as a tool for the evaluation of implementation. In M. C. Wang (Ed.), The use of direct observation to study instructional-learning behaviors in school settings. Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1974. (a) (Publication 1974/9; ERIC Document Reproduction Service No. ED 100 798)

Leinhardt, G. Evaluation of the implementation of a program of adaptive education at the second grade (1972-73). Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1974. (b) (Publication 1974/17)

Lindvall, C. M., & Cox, R. Evaluation as a tool in curriculum development: The IPI evaluation program (AERA Monograph Series on Curriculum Evaluation 5). Chicago: Rand McNally, 1970.

Mayeske, G. W. Teacher attributes and school achievement. In Office of Education (Ed.), Do teachers make a difference? Washington, D. C.: U. S. Government Printing Office, 1970. (HE5.358:58042)

Mood A. Partitioning variance in multiple regression analysis as a tool for developing learning models. American Educational Research Journal, 1971, 8, 191-202.

Raven, J. C. Coloured Progressive Matrices (Sets A, A_B, B). London: H. K. Lewis and Co., Ltd., 1956.

Rosenshine, B. Evaluation of instruction. Review of Educational Research, 1970, 40, 279-330.

Scriven, M. The methodology of evaluation. In R. Tyler, R. Gagné, & M. Scriven (Eds.), Perspectives of curriculum evaluation (AERA Monograph Series on Curriculum Evaluation 1). Chicago: Rand McNally, 1967.

Soar, R. S., & Soar, R. M. An empirical analysis of selected Follow Through programs: An example of a process approach to evaluation. The Seventy-first Yearbook of the National Society for the Study of Education, 1972, 71, 229-259.

Thorndike, R., Hagen, E., & Lorge, I. Cognitive Abilities Test (Primary I/Form I). Boston: Houghton Mifflin, 1968.

Tukey, J. W. Exploratory data analysis: Limited preliminary edition (3 vols.). Reading, Mass.: Addison Wesley, 1970.

APPENDIX A

Description of the Instructional Setting

APPENDIX A

Description of the Instructional Setting

The basic characteristics of the LRDC instructional model are that (a) it provides an environment that is adaptive to the educational needs of the student, (b) the curricula are organized and presented in a way that attempts to teach and reinforce basic cognitive skills, and (c) the student directs and controls his or her own learning within the context of the curriculum. Operationally, this means that specific subject-matter areas are broken down into units and objectives which are hierarchically sequenced in a curriculum. Placement tests, curriculum-embedded tests (CETs), and posttests are constructed for each unit so that students can be initially placed in the curriculum and so that their success in learning the material can be continuously monitored. Using the information provided by these tests, prescriptions are written for each student on a daily, weekly, or monthly basis. If a student does not pass a CET, (s)he continues working on the same objective or unit, with additional practice work. If a student does not pass a posttest, (s)he is recycled through the appropriate sections of the curriculum and retested.

The classroom normally has one teacher and one assistant or aide for 25 children. Usually, the teacher and assistant both circulate (travel) around the room during the work period; sometimes one will circulate while the other administers or corrects tests or tutors small groups. While circulating, the teacher or assistant corrects the work being done, occasionally alters prescriptions, offers brief tutorial assistance, and supports the student emotionally and academically. In general, the morning is divided into a work period in which assignments are prescribed and an exploratory period during which the children investigate curriculum-related material on their own. Classrooms vary in the degree to which students control the decision points in their daily learning situation, such as: which subject to study and when, which assignment to do first, when and when not to work on formal curriculum, when to take a posttest, when to change units, and so on. Thus, while the curriculum is relatively consistent from classroom to classroom, there are many areas in which there can be differences in implementing the program.

APPENDIX B

Description of Process Variables ... 31

Questionnaire 48

Videotape Coding Scheme 53

Videotape Rating Scheme 56

APPENDIX B

Description of Process Variables

Instruments

Thirty-three of the variables were obtained from the 15-item questionnaire which was administered twice during the year. The local educational specialist completed the forms in November and March, and the data from both administrations were then averaged.

Four variables were obtained from the student profile sheet which was cumulative and provided accurate and easily accessible information about student placement in and progress through the math and reading curricula. These sheets were kept by the classroom teacher or aide and were sent to LRDC four times during the school year.

One variable, time estimate, was obtained by using school records to determine the number of days in the academic year, questionnaire information on number of minutes in the school day, and percentage of children present on the sampled day as an estimator of attendance. These three numbers were then multiplied to yield an estimate of number of minutes spent in school during the year.

Thirty variables were obtained from videotapes recorded for 15 minutes in reading and 15 minutes in math in the fall and again in the spring and then summing over the two collections. The tapes were made by local specialists at the sites while the class was in progress. Codings were adjusted for variation in time so that all data would be based on a total of 60 minutes of videotape per classroom.

The questionnaire and videotapes focused on gathering information from seven major domains. These domains will be described briefly and the data summarized. It should be mentioned that the assignment of variables to domains is to some extent arbitrary, since one variable may, in reality, belong to several domains. Thus, the sets should not be interpreted as rigid definitions of the domains; assigning variables to domains is a convenient way of organizing and considering a rather large number of classroom process variables. Table B-1 shows the domains and variables. Tables B-2 through B-8 present the results by site.

Domains and Variables

Classroom context. This domain consists of five variables: the teacher's total years of teaching experience, the teacher's years of experience in working with the LRDC program, the total number of children enrolled in the class, the occurrence of an external event which had a severe effect (crisis) on the classroom, and the availability of a trained alternate in case of teacher absence. Although these variables do not relate directly to the LRDC program, they are likely to affect the results of that program.

Table B-2 presents the site averages for these five variables. The sites vary tremendously with regard to some of these variables. For example, teacher experience ranges from an average of 1.2 years at Site 10 to 18.5 years at Sites 1 and [illegible]. The second variable, teacher experience with the program, is dependent on the number of years that LRDC has been involved with each site. The third variable, the number of children enrolled, shows less variability than in the past. The 1972-73 range of site means was 20.6 to 27.0; the present data ranges from 20.1 to 25.8. No doubt this is a reflection of reduced enrollment. Such a reduction may affect the implementation of an individualized program in that the quality and quantity of time a teacher is able to spend with each student may be affected. Only two sites reported any crisis situation which may have had an effect on all or some of the students. The last variable in Table B-2, whether or not a trained alternate was available in case of teacher absence, shows that about

Table B-1

Classroom Process Variables

Classroom context

Total teacher experience in number of years
Total teacher experience with the LRDC instructional model in number of years
Number of children enrolled in the class
Whether or not some external event had a detrimental effect on the classroom (e.g., flood, epidemic, etc.)
Whether or not a trained alternate was available in case of teacher absence

Allocation of time

Percentage of time spent in individualized and small group activities
Percentage of time spent in large group activities
Percentage of time spent in free-choice activities
Number of minutes per day in IPI reading
Number of minutes per day in IPI math and math drill
Number of minutes per day spent watching educational TV
Time estimate-average number of minutes spent in school during the academic year

Assignment procedures and monitoring student progress

Percentage of unique assignment--the percentage of the total assignments given which were different from any other assignment given
Percentage of blocked assignments--where a blocked assignment was three or more consecutive pages or boxes with less than an equal number of isolated pages or boxes
Whether or not student's work is checked while traveling
Number of days since the last posttest was given
Percentage of students passing the last posttest
Number of days since the last test of any type was given
Initial student placement in the math curriculum
Student progress in the math curriculum
Initial student placement in the reading curriculum
Student progress in the reading curriculum

Provisions for student self-direction

Children decide what unit of math to study
Children request special tutoring
Children begin work as soon as they arrive at school
Peer tutoring occurs
Children get their own assignments
Children get their own prescription tickets
Children decide when to take a test
Children move back and forth between work and free activity

Attendance

Percentage of children present on sample days

Management procedures

Number of adults who travel during math
Number of adults who travel during reading

continued

Table B-1

Frequency of instructional event (All of the information below is available for both math and reading.)

Total number of interactions by the teacher and a child during 60 minutes--contacts
Number of times a teacher goes to a child but makes no verbal comment--checkoffs
Number of times the teacher makes a management statement or question (i.e., non-cognitive, no information is passed)--TMSQ
Number of times the teacher makes a cognitive statement which requires no response from the student--TCS
Number of times the teacher makes a cognitive question (any statement which requires an overt response from the student)--TCQ
Percentage of times the teacher makes a statement or question which is negative or punitive in nature
Number of times the teacher spends more than one minute with any one student--tutor
Number of times the teacher makes a cognitive error which is not corrected
Number of times a student initiates an interaction with the teacher--child contacts

Ratings of instructional events (1-5; 5 is high)

Active responses are sought by the teacher
The teacher uses backward chaining
The teacher models student responses
The teacher refers to earlier curricular information
The teacher focuses the attention of the students on task
The teacher solicits the opinion of students as to correctness of their responses
The teacher refers to earlier success of students
The teacher solicits the attitude of students toward the curriculum
The teacher uses praise which is contingent on student work
The teacher uses general praise of student behavior
The teacher's tutorials are clear, correct, and understandable
The teacher is able to manage the class as a whole

Table B - 2

Site Average for Basic Classroom Context Information

	Site 1		Site 2		Site 3		Site 4		Site 5		Site 6		Site 7		Site 8		Site 9		Site 10	
	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*
Total teacher *experience*	18.5	17.7	3.3	1.7	3.0	1.7	3.3	3.0	3.5	2.9	18.5	14.6	13.3	14.6	9.6	11.0	17.0	13.5	1.2	1.3
Teacher experience with the program	7.5	3.5	1.5	1.7	1.7	.5	.9	.9	1.5	.8	3.2	.4	2.8	1.5	1.1	.8	.8	.8	.2	.4
Number of children enrolled	24.0	0	21.8	.5	24.7	.8	20.1	2.7	23.0	3.7	25.8	1.9	20.5	2.4	21.3	2.1	22.2	2.3	20.4	2.4
Crisis	0	0	0	0	0	0	1.0	0	.2	.4	0	0	0	0	0	0	0	0	0	0
Trained *alternates* available	1	0	.3	.5	0	0	1.0	0	1.0	0	1.0	0	.8	.4	0	0	0	0	0	0

half of the second-grade classrooms do have an available and trained substitute.

Allocation of time. The data gathered in 1971-72 indicated that there was a positive and significant relationship between the amount of time spent in mathematics per day and the level of mathematics achievement (Leinhardt, 1974a). During 1972-73, more extensive information was gathered on the way in which individual teachers and sites spent the time available to them (Leinhardt, 1974b). The focus of this domain was centered in three areas: First, what percentage of the day is spent in specific modes of instruction--individualized and small group activities, large group activities, and free-choice activities? Second, what amount of time per day was a child exposed to specific subjects such as reading, math, or educational TV? Third, a general time estimate of total minutes spent in school.

The first three variables in Table B-3 deal with the percentage of time spent in individual, large group, and free-choice activities. We expect that at least half of the school day is spent in individualized activity, since the majority of LRDC's instructional program is individualized. From Table B-3, we note that seven of the ten sites do, in fact, spend the better part of the day in individual and small group activities. When the means of Variables 1 and 3 are combined, we find that all sites spend more time engaging in activities classified as individualized than they do in large group activities, although there are considerable differences across the sites. This is an encouraging sign since this information refers to the entire school day, not just the 3 1/2 hours during which the program is implemented. The amount of time spent in mathematics has increased for all but Site 9 since last year, whereas the amount of time spent in reading has decreased for all but Site 4. We also collected data on the amount of time spent watching educational television to determine how many of the sites use this media as a regular part of the academic day. It appears that the classrooms that do watch educational TV do so on a

Table B - 3
Site Averages on Time Allocation

	Site 1		Site 2		Site 3		Site 4		Site 5		Site 6		Site 7		Site 8		Site 9		Site 10	
	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD
Percent of time spent in individualized and small group activities	56.0	0	69.0	8.0	51.3	3.7	58.6	7.0	58.2	5.0	47.2	2.9	43.8	4.7	54.5	3.9	42.7	3.1	51.4	6.5
Percent of time spent in large group activities	31.0	0	22.3	4.6	39.7	3.2	32.9	7.0	33.2	5.7	49.3	1.9	44.8	4.7	32.1	4.1	49.0	2.6	38.6	7.3
Percent of time in free choice activities	14.0	0	8.8	6.4	9.3	1.8	12.3	5.1	9.8	4.6	4.3	2.2	12.3	3.1	14.3	4.6	8.3	1.6	9.6	3.3
Number of minutes a day in reading	55.0	0	92.5	18.9	92.5	6.1	90.0	0	84.3	9.5	98.3	2.6	81.2	7.9	86.4	7.8	110.6	8.1	115.0	7.1
Number of minutes a day in math	55.0	0	86.3	26.3	70.7	6.3	58.6	7.2	69.5	16.9	60.2	5.5	62.5	7.3	65.4	4.4	67.4	5.2	65.0	7.1
Number of minutes a day watching educational TV	10.0	0	10.5	13.3	18.8	17.6	0	0	13.5	3.5	1.7	2.9	8.3	9.3	0	0	0	0	10.6	4.0
Time estimate	57096	0	56606.8	2565.6	55242	3165.5	63997.7	1801.4	60696	1907.7	59151.5	4391.0	69240	3359.1	54075.9	1969.5	67308	2748.3	48288.2	2569.4

small scale and, therefore, little effect on achievement is expected. More sites are using educational television, however, and for longer periods of time than last year, with the most significant increase at Site 3. The last variable reported in Table B-3 is an estimate of the average number of minutes a child spends in school during the year. From the range of means it can be seen that a child may spend as little as 48,000 to as much as 70,000 minutes in school.

Assignment and monitoring procedures. This domain deals with the assignment and monitoring procedures used in a classroom. It is through the accurate matching of student needs and curriculum content that much of individualization is achieved. The first variable in this domain is the percentage of unique assignments, that is, the percentage of assignments given which are different from any other assignment given with respect to level and skill. The second variable is the percentage of blocked assignments which is defined as three or more consecutive tasks or pages assigned with less than an equal number of isolated tasks or pages. These variables tap the way in which the teacher makes assignments in class by using mathematics assignments as the sample of assignment practices in general. It appears that the teachers in all of the sites make unique assignments. The lower limit of the range of means across sites has been raised to 70.3% over the previous year's 57% (Leinhardt, 1974b). All but two sites have increased the percentage of unique assignments, and those two had minimal decreases. The percentage of assignments which are blocked decreased by about 12 percentage points across sites. The third variable relates to the kind of feedback a student gets on progress, that is, is his/her work checked during class while (s)he is at work. Table B-4 shows that at all but one site a student's work is checked during class by the teacher or aide who travels.

The LRDC instructional model also provides for frequent systematic assessment of the students' learning progress; that is, it provides procedures

Table B - 4

Site Averages on Assignment and Monitoring Procedures

	Site 1		Site 2		Site 3		Site 4		Site 5		Site 6		Site 7		Site 8		Site 9		Site 10	
	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*
Percentage of unique assignments	85.0	2.8	70.3	21.3	85.5	4.2	85.6	8.5	80.0	9.0	84.2	2.9	81.2	10.8	83.6	5.2	86.3	5.0	77.8	17.7
Percentage of blocked assignments	37.0	2.8	70.5	19.9	71.0	15.5	58.4	17.5	54.8	19.0	73.5	9.2	40.0	23.6	63.4	11.2	70.9	10.0	56.8	22.5
Work checked while traveling	1.0	0	1.0	0	1.0	0	1.0	0	1.0	0	.7	.5	1.0	0	0	0	.8	.4	1.0	0
Number of days since last posttest	8.5	3.5	--	--	17.7	5.2	16.1	4.7	16.8	2.3	16.2	2.9	23.2	8.1	14.6	4.6	16.7	5.5	18.0	9.9
Percentage passing posttest	90.0	2.8	--	--	91.0	9.3	91.3	16.6	97.0	5.3	95.3	9.6	96.8	3.6	95.3	8.9	84.4	15.6	88.2	12.5
Number of days since last test of any kind	2.3	.6	6.8	4.3	4.2	1.3	2.5	1.1	2.6	.7	3.6	.9	2.9	1.1	4.3	1.5	2.7	1.2	2.7	1.2
Student placement math	25.5	.7	--	--	11.2	1.5	10.3	2.1	11.3	1.9	19.2	2.0	9.8	.8	8.4	2.1	11.3	1.7	3.6	1.5
Student progress math	11.5	2.1	--	--	9.0	3.6	11.9	1.9	10.0	2.4	9.0	1.3	11.0	2.6	13.1	3.4	10.9	1.1	6.8	1.3
Student placement reading	14.0	0	--	--	3.0	1.7	4.3	.8	4.0	1.8	9.2	.8	6.5	2.3	6.4	1.3	4.4	.9	5.8	3.0
Student progress reading	10.0	1.4	--	--	6.0	1.7	7.6	1.5	6.3	2.3	7.7	.5	7.5	1.8	8.1	2.0	9.4	1.7	8.4	1.3

for the teacher to monitor the movement of students as they work. The assessments are generally in the form of individual, self-administered, criterion-referenced tests that are of three main types: pretests, curriculum-embedded tests (CETs), and posttests. It is expected that children will take at least one test of some type at least once every five to seven days and a posttest once every twenty days. The number of days since the last posttest varies considerably across sites. The averages range from 8.5 to 23.2. Without these two extremes, the range is between 14.6 and 18.0. The percentage of students passing the posttest is high across sites (84.4% to 97%). The next variable is number of days since the last test of any kind. The range is from 2.3 to 6.8, with the highest concentration of averages around 2.5. The last four variables in this domain deal with where the student started work in the math and reading curricula (placement) and the amount of work the student completed in those curricula (progress). It is evident from the site averages that there is a large amount of variation among sites in the placement of students in the math curriculum (3.6 to 25.5) and slightly less variation in reading placement (3.0 to 14.0). Both variations reflect basic differences in entering ability among the students at different sites. Progress within the math and reading curricula also varies (6.8 to 13.1 and 6.0 to 10.0), but there is not nearly as much variation as in initial placement.

Provisions for student autonomy (self-direction). One of the continuing long-range goals of the LRDC model is to develop independent learners. In accordance with this goal, there are many ways in which a teacher can provide opportunities within the classroom for the student to learn to exercise autonomy. Table B-5 presents the site averages on eight autonomy measures.

Attendance. Student attendance is an important variable in the implementation of the LRDC program. Generally, if the students are not present

Table B.-5

Provisions for Student Autonomy

	Site 1		Site 2		Site 3		Site 4		Site 5		Site 6		Site 7		Site 8		Site 9		Site 10	
	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*	$\overline{X}$	*SD*
Children decide what unit of math to study	0	0	--[a]	--	0	0	0	0	0	0	0	0	.67	.82	0	0	0	0	0	0
Children request special tutoring	0	0	--	--	.67	.52	2.0	0	1.17	.75	1.17	.98	2.0	0	1.63	.52	1.11	.33	.4	.55
Children begin work as soon as they come in	0	0	--	--	.83	.98	.71	.76	1.0	.89	.33	.52	.83	.98	.50	.76	.22	.44	.8	.45
Peer tutoring occurs	0	0	--	--	1.67	.52	1.57	.54	1.5	.55	1.83	.41	1.83	.41	2.0	0	1.22	.44	1.0	0
Children get own assignments	1.0	0	--	--	.17	.41	.43	.54	1.83	.41	1.17	.98	0	0	.63	.92	1.44	.88	.2	.45
Children get own prescription ticket	0	0	--	--	.5	.55	.14	.38	0	0	.83	.41	.17	.41	0	0	.11	.33	.2	.45
Children decide when to take a test	0	0	--	--	.33	.52	1.14	.69	.67	.82	1.0	.89	.67	.82	.38	.52	0	0	.4	.55
Children move back and forth between work and free activity	0	0	--	--	0	0	0	0	.67	.82	0	0	0	0	0	0	0	0	0	0

[a]Information not available.

in school, they cannot benefit from instruction. Although an individual program makes student absenteeism less damaging than a traditional program, it is still an important variable to take into account. Good attendance can also be taken as an indication of a positive attitude toward school in general. Table B-6 reports student attendance based on the percent of children present on two randomly selected days. The means in general are high and comparable to last year's data.

Management procedures. The number of adults who travel during mathematics and reading is reported in Table B-7. It appears that most classrooms have two adults traveling during these classes. This is contrary to expectations, and it is not clear that this is the most efficient use of teacher time.

Instructional event. This domain deals with ways in which teachers may differ in in-class behaviors. Our concern here is with the basic interactions between a teacher and a student during a 60-minute period. The variables included relate to: the rate at which the teacher moves from child to child, the type of contact (s)he makes with the child, the length of time (s)he spends with any one child, the effect of the interaction (i. e., positive or negative), the number of cognitive errors the teacher makes, how often the students initiate an interaction with the teacher, and a list of teacher characteristics on which (s)he is rated. Teachers vary along all of these dimensions, and an individualized program may require, for example, a faster rate of movement and greater frequency of contacts than a traditional one. Through videotaping classes in progress, we hoped to be able to take a closer look at the variables deemed significant in the domain of the instructional event.

Table B-8 presents the 30 variables derived from videotaping teacher-child interactions in each second-grade classroom for 30 minutes of reading and 30 minutes of mathematics. The total number of contacts in reading

Table B - 6

Student Attendance

	Site 1		Site 2		Site 3		Site 4		Site 5		Site 6		Site 7		Site 8		Site 9		Site 10	
	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD
Percent of children present	96.0	0	94.3	4.3	94.2	3.8	96.7	0.1	95.0	3.1	95.8	3.4	94.8	3.4	95.9	3.8	96.3	2.4	80.4	4.3

Table B - 7

Management Procedures

	Site 1		Site 2		Site 3		Site 4		Site 5		Site 6		Site 7		Site 8		Site 9		Site 10	
	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD
Number of adults traveling math	2.0	0	1.8	.5	2.0	0	2.0	0	1.8	.4	2.2	.4	1.5	.5	1.9	.4	2.0	0	2.4	.6
Number of adults traveling reading	2.0	0	1.8	.5	2.0	0	1.1	.4	1.7	.5	2.5	.8	1.5	.5	1.9	.4	1.8	.4	2.4	.6

Table B-8

Site Averages on Instructional Event

	Site 1		Site 2		Site 3		Site 4		Site 5		Site 6		Site 7		Site 8		Site 9		Site 10	
	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD
Total reading contacts	80.0	62.2	--[a]	--	45.5	9.1	35.6	19.0	41.5	11.0	40.7	19.9	37.8	12.4	46.1	21.3	33.2	13.8	48.0	17.6
Reading checkoffs	2.5	2.1	--	--	.2	.4	.3	.5	.5	1.2	0	0	1.2	1.0	.1	.4	0	0	0	0
Reading TMSQ	37.5	34.6	--	--	22.8	7.8	18.6	15.6	24.2	12.0	11.7	6.7	11.5	5.5	23.1	17.6	14.0	11.3	40.8	18.7
Reading TCS	6.0	8.5	--	--	.7	1.6	1.3	1.9	1.0	1.3	1.8	1.8	1.0	1.3	1.6	2.3	.3	.5	.8	1.3
Reading TCQ	29.0	12.7	--	--	27.5	5.2	15.9	4.7	16.3	6.3	[illegible]	18.2	25.5	5.0	23.6	5.9	21.3	7.6	9.4	3.6
Reading % negative	2.5	3.5	--	--	.3	.8	1.0	2.2	1.7	3.2	0	0	1.0	1.5	0	0	1.0	2.0	4.6	6.2
Reading tutors	.5	.7	--	--	4.8	2.4	8.6	3.9	4.2	2.5	5.3	3.6	5.8	2.3	6.6	2.8	7.8	3.1	6.4	2.2
Reading errors	.5	.7	--	--	1.0	.9	.3	.5	.8	1.6	0	0	.3	.5	.1	.4	.4	.9	0	0
Reading child contacts	22.0	22.6	--	--	5.0	3.6	11.0	8.5	3.3	3.6	7.0	5.1	9.5	5.1	12.3	5.5	7.4	3.2	17.2	7.3
Math total contacts	82.0	70.7	--	--	58.0	35.1	44.4	14.3	40.2	11.1	29.8	11.5	43.8	14.6	55.8	25.5	51.8	18.3	45.2	11.4
Math checkoffs	2.0	2.8	--	--	1.3	1.9	0	0	1.7	3.6	0	0	.8	2.0	.9	.8	.2	.4	1.6	2.6
Math TMSQ	42.0	33.9	--	--	22.0	7.3	22.3	12.0	15.5	7.2	10.7	12.6	17.8	12.1	18.4	8.4	21.9	12.0	28.6	5.0
Math TCS	19.0	24.0	--	--	18.2	25.6	5.7	3.6	6.7	2.6	1.8	1.7	1.3	1.5	8.6	8.6	7.1	4.2	3.0	2.5
Math TCQ	17.0	4.2	--	--	21.8	10.1	20.4	2.4	19.0	6.7	10.7	4.6	24.2	4.7	28.9	12.5	25.0	8.0	13.4	3.4
Math % negative	1.0	1.4	--	--	3.3	2.8	1.4	3.0	2.3	3.2	.8	2.0	1.0	1.7	.8	1.4	1.3	3.0	4.8	7.2

continued

Table B-8

	Site 1		Site 2		Site 3		Site 4		Site 5		Site 6		Site 7		Site 8		Site 9		Site 10	
	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD
Math tutors	3.0	4.2	--	--	3.8	4.1	6.4	4.8	4.7	1.5	7.2	2.6	4.5	2.8	6.5	4.4	4.4	2.7	8.2	2.0
Math errors	0	0	--	--	1.2	.8	.3	.5	0	0	0	0	.5	.8	.1	.4	.4	1.3	.4	.9
Math child contacts	19.0	12.7	--	--	4.7	1.6	13.1	7.2	6.2	6.6	5.7	5.5	11.7	7.2	10.3	4.8	12.4	5.1	9.0	4.0
Active response sought	2.5	.7	--	--	4.0	.9	4.3	.8	3.3	.5	4.2	.4	4.0	.6	3.9	.4	3.9	.6	3.2	.8
Backward chains	1.0	0	--	--	1.2	.4	1.4	.5	1.3	.5	1.3	.5	1.3	.5	1.6	.5	1.2	.4	1.6	.5
Models responses	1.5	.7	--	--	3.0	.9	2.4	.5	1.7	.5	2.5	1.0	2.7	1.0	3.1	.8	2.8	.8	2.2	.4
Refers to earlier info.	1.5	.7	--	--	1.7	.8	1.9	.7	1.2	.4	1.3	.5	2.2	.8	2.0	.8	1.7	.7	2.8	1.1
Focuses attention	2.5	.7	--	--	3.8	.4	3.9	.4	3.0	.6	4.3	.5	4.3	.5	4.1	.6	3.9	.3	3.0	0
Solicits opinion on correctness	2.0	0	--	--	1.8	.4	2.1	.4	1.5	.8	2.0	.9	1.5	.8	1.9	.4	2.3	.7	2.8	.8
Refers to earlier success	1.5	.7	--	--	1.5	.8	1.3	.5	1.7	.8	1.5	.5	2.0	.6	1.6	.5	1.4	.5	1.2	.4
Solicits attitude toward curriculum	1.0	0	--	--	1.5	.5	1.1	.4	1.3	.8	1.5	.5	1.5	.5	1.8	.9	1.7	.9	1.4	.5
Contingent praise	2.5	.7	--	--	3.8	.4	4.3	1.0	3.3	.8	3.8	1.0	3.7	1.0	4.1	.6	4.1	.6	2.2	.4
General praise	2.0	1.4	--	--	2.8	.8	3.0	.8	2.2	1.2	2.8	1.2	2.8	1.0	3.3	.5	3.2	.7	1.4	.5
Tutoring	2.5	.7	--	--	3.5	1.0	2.9	.4	3.0	.6	3.8	.8	3.0	.9	3.4	.5	3.6	.9	2.6	.5
Management	3.0	1.4	--	--	3.3	.8	3.6	.5	3.0	1.1	4.5	.8	3.8	1.2	4.3	.5	3.9	1.4	1.6	.5

[a] Information not available.

and in mathematics shows that there is a great deal of difference across and within sites in the rate at which the teacher travels. The respective ranges of means are from 33.2 to 80 and 29.8 to 82, with overall means of 45.38 and 50.11.

The next four variables in reading and math deal with the types of contact the teacher has with the child. The teachers appear to make most of their contacts in management areas and in asking cognitive questions. Teachers appear to make few cognitive statements in reading. In mathematics, the number of cognitive statements is greater, but means are less than those for cognitive questions with the exception of Site 1.

The percentage of negative statements is small in both reading and math. The tutors variable is characterized as any interaction between the teacher and a child which lasts longer than one minute. Here again, the frequency is relatively small, with mean reading tutors ranging from .5 to 8.6 and math tutors ranging from 3.0 to 8.2. The interpretation of this variable is ambiguous at this time. That is, do tutorial interactions increase the quality of instruction with a given child to the extent that it overcomes the diminished rate of travel and frequency of contacts? Or, is it more valuable to interact with more children, even if for a limited time? The number of cognitive errors made by the teacher in reading and math are very minimal (0 - 1.2).

The final variable directly related to reading and math was the number of times the child initiated the interaction with the teacher by either making a statement or asking a question. Once again there is substantial variation across sites (means of 3.3 - 22.0 in reading and 4.7 - 19.0 in math), as well as within sites (standard deviations of 3.2 - 22.6 in reading and 1.6 - 12.7 in math).

The last 12 variables were scored on a rating schedule from 1 to 5 based on observations of the classroom videotapes, with 5 meaning the teacher used the named skill often and well. Teachers in all sites appear to seek active responses from the child, since the lowest mean is 2.5 (only one site) and the others range from 3.2 to 4.3. On the other hand, it would appear that there is very little backward chaining in instruction, since all means are between 1.0 and 1.6. "Modeling responses" seems to be a more frequent form of instruction, but is still not extensive across sites, and referring to earlier curricular information occurs only slightly more than backward chaining.

Table B-8 also indicates that teachers across sites do focus the attention of students on task. This may be related to seeking active responses, since the means for both variables are quite similar and of the same magnitude. The next three variables seem to occur only occasionally with smaller variation across sites. Two of these variables relate to soliciting feedback from the child on specific work done and on the curriculum in general. Contingent praise refers to positive reinforcement for specific work completed, whereas general praise is not related to any specific cognitive task. Across sites, contingent praise is used more often than general praise and appears to be used frequently.

The next two variables relate to tutoring and management behaviors. The tutoring rating was based on how well the teacher tutors a child: Is it clear, concise, and understandable by the child? Management was rated according to how well the class as a whole functions: Do the children know what to do and do it, and are the class and the teacher moving efficiently? The program seems to be working well on both counts, since the ratings in these two areas are fairly high with the exception of Site 10 in management.

Questionnaire

Name of Individual
Filling Out Questionnaire ______________________

Date ______________________

District ______________________

School ______________________

Teacher's Name ______________________

1. How many years of teaching experience (prior to this school year) does the teacher have? ______________________

2. How many years has the teacher been using PEP or IPI prior to this school year? ______________________

3. How many children are enrolled in this classroom? ______________________

4. How many children are present today in this room? ______________________

5. How many adults are normally in the room? ______________________

6. How many adults normally travel at any one time?

   During math: ______________________

   During reading: ______________________

7. Does the teacher (or aide) check the majority of the children's IPI/PEP work while traveling?

   For math work: ______________________

   For reading work: ______________________

   If either of the above is no, when is the majority of the work checked?

   ____

   ____

DATE THIS QUESTION IS BEING FILLED OUT ________________

8. Please list each of today's IPI/PEP math and/or quantification assignments for each student. Please also list the DATE and SCORE of the last POSTTEST given. Record the posttest score regardless of whether or not it was passed. (Note: If assignments are made by the week, record the entire assignment.)

	Unit	Level	Skill	Pages	Please record the date and score of the last POSTTEST given.
Example:	Add/Sub	B	4	1-4, 7 1, 8, 10	POSTTEST 3/18/74 74%

Reminder: Did you record all of the dates and scores?

9. Check those activities which you observe occurring regularly in this room.

Students obtain own assignments for math (as opposed to having them passed out).	______
Students obtain own notebooks for reading.	______
Students sometimes write their own prescriptions (other than exploratory) for math, reading, or science.	______
Students are consulted as to the mode of instruction (tapes, books, texts, or games).	______
Students select time of day for work in subject-matter area.	______
Students help to select which unit of math (geometry, money, fractions) will be studied.	______
Students suggest when to take CET (either by requesting or delaying it).	______
Students suggest when to take <u>Unit</u> test.	______
Check only one: Students move back and forth between exploratory and prescription areas.	______
Check only one: Students go to exploratory any time after completion of predefined list of assignments.	______
Check only one: Teacher decides when exploratory period begins for the majority of students.	______
Students request or bring in activities for exploratory areas.	______
Students request small group lesson.	______
Students request special tutoring.	______
Student's opinion of curriculum or task is sought. Give an example: ______________________	______
Students sometimes choose to work with other students rather than alone.	______
Students may start work as soon as they come in in the morning (or after lunch) rather than waiting for the teacher to start the period.	______

Peer tutoring occurs. ____________

Teacher seeks information about the attitudes of students toward their work. ____________

10. Please give an approximate number of hours and minutes which any one student would spend during an entire school day in each of the following instructional situations. The total number of hours and minutes should add up to approximately one whole day, not just the IPI part of the day.

____________ Student works alone with occasional adult or peer tutoring.

____________ Student works in a small group with adult instruction (2-8 children).

____________ Student works in a large group with adult instruction (9-whole class).

____________ Student is engaged in free-choice activity.

____________ Other. Specify ____________________________

11. Please list the number of minutes per day students in the classroom spend in the following activities:

____________ IPI Reading

____________ IPI Math

____________ Some form of math drill

____________ Wathing educational TV
(Note: If educational TV is unavailable in your school due to either lack of reception or availability of TV monitors, please indicate above.)

12. Is a math maintenance program being used in this room (systematic drill of number or algorithm facts three or more times a week for 10 minutes or more)? ____________

13. Any unusual events which in your opinion would strongly affect the nature of this classroom (such as floods, severe illnesses, etc.)?

____

____

14. Does this classroom have trained, permanent alternates who substitute in the case of teacher or aide absence? ____________

15. Observe the teacher for 10 minutes during the math or reading period and record the number of negative comments made by her or him. (A negative comment is one that generally orders a child to stop some activity or is sarcastic or punitive toward a child or group of children.)

Videotape Coding Scheme

Videotapes are observed using the teacher-student interaction as the main unit of coding. Every time the teacher changes from one child to another, that interaction is recorded according to what the teacher says and the response(s) of the child or children. There are eleven categories that may be checked for each interaction. The following is a summary of the guidelines used to determine if a category should be checked.

1. Checkoff (✓)

   The teacher checks a child's work and leaves without making any comment to the child, i.e., there is no verbal communication.

2. Teacher Management Statements or Questions (TMSQ)

   The teacher says something to the child (or children) which contains no cognitive information but serves a managerial function only. Management statements or questions may concern curricular materials, discipline, personal functions, or information about placement within a curriculum. General examples are: "Yes, you may go to the bathroom"; "Get your materials ready"; "Has the aide checked your work?"; "Be quiet; sit down"; "Go to page 10"; "Good" or "Right" (if "good" and "right" occur by themselves, they are considered management information); etc.

3. Teacher Cognitive Statements (TCS)

   The teacher makes a statement containing cognitive information which does not elicit a response from the child or which elicits a response which is an echo of the teacher's statement. Cognitive statements may occur when the teacher reads directions; the teacher asks a rhetorical question; the teacher asks a question and then answers it; the teacher gives the command, "Repeat 2 + 2 = 4," and the child says "2 + 2 = 4"; or the teacher lectures.

4. Teacher Cognitive Questions (TCQ)

   The teacher asks a question about cognitive material which elicits a response from the child. Examples of teacher cognitive questions are: "How much is 2 + 2?"; "Read me the sentence that tells what Sam is doing"; "Read the directions which tell you what to do on the next page"; and "Read this sentence." It can also be the kind of statement which appears to be managerial but which clearly contains cognitive information. Consider the following set of questions:

"How many pages did you do today? How many were you supposed to do? How many are you going to do tomorrow? Did you do more or less than you had planned to do?" Questions that give detailed discussion about the titles of books, how to spell them, and how to map them out are also considered to be cognitive. Although the above examples are teaching children skills not normally taught in the curriculum, they are clearly cognitively based skills and are, therefore, counted as cognitive questions or statements. If a cognitive question occurs with a cognitive statement, only the cognitive question is counted in the summary of the videotape data.

5. Child Cognitive Statements (CCS)

The child makes a statement regarding cognitive information in response to a question, without a prompt, or reads some cognitive material--a story, directions, and so on. (This category is only coded in the presence of a teacher because the videotape camera follows the teacher.)

6. Child Cognitive Questions (CCQ)

The child asks a question which elicits a response from the teacher. This may be in the form of a statement or a question. For example, "How do I find this answer?" or "I don't know how to do this."

7. Child Management Questions (CMQ)

The child asks a question or makes a statement which elicits a teacher response dealing with a management area. Examples are: "Where is the tape for this lesson?"; "I don't have a pencil"; "May I sit next to Susan?"; etc.

8. Child Management Statements (CMS)

The child makes a statement requiring no response from the teacher in a management area. Examples are: "I fell on my way to school"; "Johnny hit me"; "The aide checked my paper"; etc. The rules of statements versus questions are not semantic; they are based on whether or not a response was made. (The same rules apply with Teacher Statements.)

9. Tutor

An interaction is counted a tutor if it lasts over one minute.

10. Error

The teacher makes an error and does not correct it; for example, "2 + 2 = 5." This is not a pedagogical error; this is an error of information.

11. Negative

Any category in which the teacher's statement is essentially punitive is considered to be negative. The statement can be considered negative when the tonal quality of what is said is punitive, not just the content of the statement. "You did very well on that work," if said in a sarcastic and degrading manner, may mean that the teacher is telling the child how terrible (s)he is doing. Most negative statements involve demanding the stopping of behaviors on the part of the child. Examples of this are: "Stop it"; "Sit down"; "You weren't listening". "No" or "not" might be the cue. Repetition such as, "I've told you that three or four times," or a negative facet attributed to the child's overall performance, as "You never listen when you are being told something," are also cues to negative statements.

Videotape Rating Scheme

At the end of observing the videotape of a teacher, a rating scale is filled out. There are 12 rating categories. Ratings are used instead of frequency counts because frequency coding did not turn out to provide additional information in these areas, and it was easier to rate than to count frequencies for these categories. The coders are instructed to know the 12 categories before viewing the tape. They are not to fill out the rating until they have watched the whole tape, but they may take notes. Each category is rated on a scale from 1, indicating never, to 5, indicating frequently.

1. Active Responses Sought

   The teacher elicits active responses from the students. That is, students are required to verbalize, manipulate, write, or perform some active form of behavior when the teacher interacts with them, as opposed to passive behavior such as listening.

2. Teacher Backward Chains an Explanation

   The teacher, while tutoring, gives the child a complete explanation and elicits behaviors from the child in a backward chain. The teacher uses the last stimulus of a chain as a reinforcer for the behaviors required in the next to the last step, and so on.

3. Teacher Models a Response

   The teacher very clearly shows the child what a response consists of. For example, the teacher says, "3 + 4 = 7. Write the 7 in the box." (S)he then shows the child how to do it and watches as the child does the next problem.

4. Teacher Refers to Earlier Curricular Information

   The teacher mentions, during an interaction with a student, information that was previously learned as in, for example, "You sounded out words like this before." It is a communication to the child that the task the child is facing is something not totally new.

5. Teacher Focuses Child's Attention on Task

This may be judged not only by what the teacher says or does, but also by what the child is doing. That is, does the child look off into space or at the materials? Does the teacher let the student wander off on a tangent, or does (s)he keep the student on task?

6. Teacher Solicits Child's Opinion on Correctness of Response

The teacher asks the child to determine whether his/her response is correct.

7. Teacher Refers to Earlier Success

The teacher mentions something the child has done well in the past, as, for example: "You really did well with multiplication yesterday; division is the reverse process. See if you can figure out how to do these problems. You have done this kind of thing very well in the past."

8. Teacher Solicits Information on Attitudes Toward the Curriculum

The teacher asks the child how (s)he feels about the curricular materials or information.

9. Teacher Uses Contingent Praise

The teacher praises a student for work that is well done, for example, "You added these two numbers correctly; good." It may even be telling the child that (s)he consistently did something correctly This is contrasted with the next category, general praise.

10. Teacher Uses General Praise

The teacher praises a student in a nonspecific way: "You are working well," "You have done good work," or "You are in a great frame of mind today." The child must interpret what is being praised. This does not mean that No. 9 and No. 10 are mutually exclusive. No. 9 is specific in relation to the behavior, and No. 10 is general.

11. Tutoring

This judgment is based on the clarity, conciseness, and accuracy of the teacher's tutorials. That is, is (s)he explaining things clearly, understandably, and logically? Is (s)he to the point? Does (s)he answ the child quickly and flexibly? Does the child seem to follow the expla tion and use the materials appropriately? If the student does not unde stand, does the teacher take a new approach, become more concrete, or find some innovative way of expressing the thought?

12. Class Management

This judgment is based on how well the classroom works, not necessarily on how low the noise level is. This category was developed because it was clear that despite the fact that teachers had very different management procedures, some being quite constrained and strict and others being quite loose, within each type there were very different successes. That is, there are classrooms where the teachers are very structured and want a quiet room but are not always telling their students to be quiet. Equally, there are classrooms that are open, free, and quite noisy and in which the teacher seems to be quite comfortable. What is basically being measured is the degree of harmony between what the teacher seems to want and what the children seem to be doing. Is the teacher fighting whatever is going on in the classroom? Do the children seem to know what is expected of them? Do they get materials easily? Is everyone functioning individually as well as in a unit? Is the process smooth and even, without major disruption?