DOCUMENT RESUME

ED 095 174 95 SP 008 354

AUTHOR Evertson, Carolyn M.; Brophy, Jere E.
TITLE High-Inference Behavioral Ratings as Correlates of Teaching Effectiveness.
INSTITUTION Texas Univ., Austin. Research and Development Center for Teacher Education.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 74
CONTRACT OEC-6-10-108
NOTE 35p.; Expanded version of a paper of the same name presented at the Annual Meeting of the American Educational Research Association (58th New Orleans, Louisiana, February 25-March 1, 1973)

EDRS PRICE MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS *Academic Achievement; *Effective Teaching; Performance Criteria; *Student Improvement; *Students; Success Factors; *Teacher Behavior

ABSTRACT

High-inference measures of teacher process variables were taken on a sample of 31 teachers selected because of their consistency in producing student learning gains on the Metropolitan Achievement Test and were correlated with student outcome measures. Correlations showing the strength of relationships with success in producing student gains are presented. Data represent findings from the first year of a 2-year study attempting to isolate correlations of effective teaching. (Author)

HIGH-INFERENCE BEHAVIORAL RATINGS AS CORRELATES OF TEACHING EFFECTIVENESS[1]

Carolyn M. Evertson and Jere E. Brophy[2]

The University of Texas at Austin

[1]This paper is an expanded version of a paper of the same name presented at the Annual Meeting of the American Educational Research Association, held February 25-March 1, 1973, in New Orleans, Louisiana.

This study was supported in part by National Institute of Education Contract OEC 6-10-108, Research and Development Center for Teacher Education. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that office should be inferred.

[2]The authors wish to acknowledge and thank Don Veldman for his statistical design suggestions; Shyam Ahuja, Marilyn Arnold, Jim Blackwell, Maria Buczynski, John Crawford, Susan Florence, Teresa Harris, Carol King, Karen Mays, Mark Mays, Nancy Moore, Piara Pannu, Brian Peck, Kathleen Senior, Carol Watkins, Michael Weissberg, and Andrea Winter for their help in classroom coding and data preparation; John Sheffield and Jim Sherrill for their help in data analysis; and Marilyn Arnold and Karen Mays for help in manuscript preparation.

HIGH-INFERENCE BEHAVIORAL RATINGS AS CORRELATES OF TEACHING EFFECTIVENESS

Carolyn M. Evertson and Jere E. Brophy
The University of Texas at Austin

The findings reported in this paper are based on data taken from the first year of a two-year naturalistic study of elementary teachers selected because of their consistency across 3 years in producing student learning gains on the Metropolitan Achievement Tests.

Of 275 teachers working at grades two and three, 88 second grade and 77 third grade teachers who had taught at the same grade for 5 years or more were identified. Thus the sample included only experienced teachers. Since the district administered the Metropolitan Achievement Tests each fall, it was possible to use these data for determining teacher effectiveness. Students' residual gains on the subtests of the Metropolitan battery were computed by using scores from the beginning of the year as pre-score covariates for adjusting gains between that testing and the following testing a year later at the beginning of the following grade. In addition to computing residual gains by using pre-scores as covariates, student differences were taken into account by separately computing the data for Title I vs. Non-Title I schools, for boys vs. girls, and for each of the 3 years separately. Thus, for example, the formula used in computing the residual gain scores for a boy in second grade in a Title I school in 1969 was based on the distribution of pre- and post-scores for all boys in the second grade in Title I schools in 1969.

Each student took 3 language arts subtests (word knowledge, word discrimination, and reading) and either 1 or 2 arithmetic subtests. The number of arithmetic subtests available depended upon which Metropolitan battery had been

used and on how the data were recorded in the school records. For example, a given child's arithmetic data might contain an arithmetic computation subtest only, a combination score reflecting both computation and reasoning, or 2 separate scores, 1 for computation and 1 for reasoning.

Two sets of computations of residual gain scores were made for each grade, because of differences in the test batteries used in Title I and Non-Title I schools. For each of these 4 data sets, residual gain scores were computed for each student within sex and within each of the 3 years on each subtest, using the student's pre-score as a covariate. These residual gain scores for students were then collated by classroom, and a mean residual gain score was computed for each teacher for each subtest for each of the 3 years included in the study (Brophy, 1972).

Many teachers showed constancy across subtests within years as well as within subtests across years, so that teachers who produced generally consistent gains across subtests and across the two sexes could be identified. The 31 teachers included in the process observation study the first year (1971-1972) were selected from this consistent group.

The teachers were divided fairly evenly between second and third grade, and the socio-economic status of their students ranged from upper middle to lower lower class. Several classes had a high concentration of black and Chicano students. The teacher data included process measures of classroom behavior and personality and attitude data from pencil and paper tests. These process measures included both low-inference behavioral observation systems and high-inference ratings. Each teacher was observed for 2 mornings and 2 afternoons during the spring semester, for a total of about 8 hours. The main low-inference coding instrument used was based on the Brophy-Good Dyadic Interaction

System (Brophy and Good, 1970). The system allowed for coding such variables as teacher vs. student initiation of contacts, types of interactions (academic, procedural, or behavioral-disciplinary), difficulty level of teacher questions, quality of student responses, quantity and quality of teacher feedback and evaluative reactions to student response and student work, and the teacher's method and general effectiveness in handling classroom management and disciplinary problems.

A second coding instrument (created by project staff member Nancy Moore) was used on a subsample of 10 teachers (5 high and 5 low effective) who were observed twice during group instruction activities. This instrument was especially constructed to measure group instruction methodological variables, such as lesson composition, sequence, and clarity, teacher questioning patterns, and handling of seatwork assignments (see Brophy and Evertson, 1973, for a detailed description and discussion of the findings from these systems).

Teachers were asked to fill out the pencil and paper attitude and personality measures included in the COMPASS battery developed by the Research and Development Center for Teacher Education (Veldman, 1972). These measures dealt with levels of teacher concerns about teaching (Fuller, 1969), the teacher's overall self-concept and her concept of herself as a teacher, her coping style, other aspects of her personality, and certain demographic data (see Peck and Veldman, 1973).

The high-inference measures to be reported here were chiefly of two types. The first was a set of 12 classroom observation scales developed by Emmer (1973) from factor-analytic studies of interaction variables common to several of the widely used behavioral observation systems. These 5-point rating scales included ratings of student attention, teacher enthusiasm, clarity, positive and negative affect, task orientation, cognitive level of questions, student passivity, pupil-

pupil interaction, and percentage of time spent in lectures and demonstrations. The scales were marked several times during each of the 4 observations by classroom coders who were also using the expanded Brophy-Good System discussed in Brophy and Evertson (1973).

Secondly, following their last 2 visits to each teacher's classroom, observers filled out 41 high-inference ratings and 15 high-inference checklists and percentage estimates. Items for these measures were gathered from several sources, but most of them dealt with variables which were not covered in the low-inference behavioral coding system. Inter-observer agreement on Emmer's Classroom Observation Scales ranged from 72% to 95%, with an average agreement of 83%. The ratings agreement ranged from 50% in one case to 98%, with an average agreement of 78%. The raw agreement data are presented in Tables 2 and 4 for the checklists and percentage estimates. No actual inter-observer agreement percentages were computed for these measures because scores would vary depending upon the formulas used. It was decided that the raw data give a more accurate picture of the real agreement.

During each classroom observation, coders also noted the amount of time allotted to each subject as well as to special activities such as storytelling, TV, art, and transitions. The time utilization results are reported in the first part of Table 2.

For the most part, the high-inference measures bore out the findings from the low-inference behavioral coding systems (Brophy and Evertson, 1973). The correlations reveal several interesting findings. First, with few exceptions, significant correlations between process variables and gain scores typically involved only one or two of the subtests, despite the fact that teachers had been selected because of their general consistency in producing student gains across all subtests. Thus it appears that certain teacher behaviors are more

important for student gain in some subject areas than in others.

Many variables which correlated significantly with pupil gains in other studies did so in this one, but many did not. For example, teacher warmth, cognitive level of questions, enthusiasm, amount of student talk, peer-tutoring, solidarity with students (rapport), and patience all failed to show significant correlations with gains. There remains the possibility that some of these variables are curvilinearly related to the criterion, however. Analyses addressing this question are being completed and will be summarized in a future report.

In addition to correlations across the total sample, analyses were also performed separately for Title I (low SES) and Non-Title I (high SES) schools.

Results

Data will be discussed within two general categories and grouped within categories under three broad headings: 1.) variables which pertain to non-academic classroom management, control, and organization; 2.) variables which can be considered teacher techniques or behaviors in presenting subject matter; 3.) student variables which are pupil behaviors not under the direct control of the teacher.

The first category will include variables which are correlated significantly in the total sample and, either significantly or in the same direction, in each of the two subgroups (Title I and Non-Title I).

Correlations for the Total Group

A. Classroom management and control (positive relationships)

1. Room appearance is attractive and uncrowded.
2. Students are expected to care for their own needs without getting permission.

3. Teacher explains rules or decisions when the reasons for them are not obvious.
4. Teacher is well-organized and well-prepared; she is task-oriented and doesn't waste time.
5. Teacher monitors the class regularly, knows what is going on and keeps herself aware of events in other parts of the classroom.
6. Transitions between activities are smooth and efficient with little wasted time.
7. Classroom jobs are determined by some automatic system instead of the teacher's spending a great amount of time deciding who is going to do particular tasks.
8. A high percentage of teacher structured time in math (correlates with both the reading and arithmetic computation subtests).

B. Classroom management and control (negative relationships)

1. Chaotic, poorly planned class schedules.

C. Teacher techniques or methods (positive relationships)

1. Teacher often addresses problems or questions to the whole class as opposed to individuals or subgroups.
2. Teacher spends a high percentage of time in lectures, answering pupil questions, demonstrations, and presenting substantive information to the class as opposed to questioning students, giving procedural directions, and evaluating (praising and criticizing) student behavior.
3. Seatwork assignments are appropriate to the task at hand and to the individual level of each child.
4. When a student doesn't understand his seatwork, she asks another child to explain rather than allowing interruptions in what she is doing, or delaying the child in order to explain later.
5. When helping a child she usually bends close and gets down to his level.

6. When a child is stuck on a word in reading group, she gives him the word.
7. Teacher uses visual aids, pictures, and filmstrips as materials for free time activities.

D. Teaching techniques or methods (negative relationships)

1. Teacher's allowing interruptions while she is working with individuals or group to explain to a child who doesn't understand.
2. Giving overly long, explicit, repetitive directions. (Possibly over-dwelling or needlessly rehashing information, which most of the class already knows, serves to "turn off" or lose students.)
3. Having books available in the room (not necessarily used).
4. Asking another child to give the word when a child is stuck.
5. A high percentage of teacher structured time.
6. When a child is stuck on a word, giving a context clue or definition.

E. Student behaviors (positive relationships)

1. High general level of student attention and a high percentage of students attending when lessons are explained or directions given.

F. Student behaviors (negative relationships)

1. Copying from a neighbor rather than asking for help when having trouble with an assignment.
2. When students are not working, they are being distracted by activities in other groups going on in the room.

In the second category are those variables which were significantly correlated with student learning only in one group.

Correlations in Title I

A. Classroom management and control (positive relationships)

1. Keeping child in after school as punishment.

B. Classroom management and control (negative relationships)

1. Underreacting to discipline problems so that serious problems go unresolved. (This receives some support from the behavioral data taken with the classroom coding system. Thus it seems more important for the teacher to stop firmly any control problems before they get out of hand than it is to criticize or punish offenders after the problem or behavior has gotten out of control.)
2. Allowing 4 or more students up at anytime without permission as long as they stay quiet. (It is possible that this is indicative of poor classroom control. This lax control could prevent time being spent on subject matter.)
3. Having a well-established routine which minimizes interruptions; room runs "automatically".

C. Teacher techniques or methods (positive relationships)

1. Teacher's ability to admit her own mistakes and laugh at herself or use the occasion to teach.
2. Teacher demonstrates showmanship and expressiveness.
3. Clear explanations.
4. Having available and using listening centers, aquariums, and other "looking" exhibits.
5. Giving the whole word to a child instead of a hint or unrelated clue when he is stuck during reading group.
6. Using visual aids (picture files, filmstrips) for free time activities.
7. A high percentage of teacher structured time spent in reading groups.
8. Goes to the child's desk to give help rather than having child come

up to her desk.

9. When a child doesn't understand his seatwork, teacher asks another child to explain.

D. Teacher techniques or methods (negative relationships)

1. Giving a child who is stuck during reading a context clue or definition rather than the whole word.
2. Having either instructional or non-instructional games available in class.
3. Use of concrete (candy, money) items, special jobs, or classmates clapping or cheering as rewards.
4. Teacher's identifying with the class and promoting a "we" feeling as opposed to standing aloof and separating herself from the class.
5. Accepting and integrating student ideas instead of rejecting unsolicited ideas and sticking with a preplanned format. (This finding is borne out to some degree in the behavioral data which shows that for Title I, a high proportion of student-initiated comments vs. teacher afforded public response opportunities was negatively related to student learning gains.)
6. Gives complete detailed instructions; prevents errors before they happen.
7. A high percentage of teacher structured time spent in language arts and in special activities.
8. When a child doesn't understand his seatwork, the teacher delays him then explains later, if she happens to be busy.

E. Student behaviors (positive relationships)

1. High pupil-pupil interaction which is class relevant.
2. When students are not working they are playing as opposed to engaging in some other activities such as daydreaming or disturbing others.

F. Student behaviors (negative relationships)

1. Student obedience and compliance. (Possibly this variable is actually a measure of teacher punitiveness which does show a negative relationship in some of the other data.)
2. Student daydreaming or getting materials for free time activities when not working.

Correlations in Non-Title I

A. Classroom management and control (positive relationships)

1. Punishments for misbehavers involve discussing the incident with the child without scolding as opposed to a more physical means of control.
2. Before starting a lesson or explanation, teacher says nothing and waits for the class to quiet down rather than trying to talk over the noise or signaling with a bell or clicker.

B. Classroom management and control (negative relationships)

1. "Busy", cluttered classrooms.
2. Boring, repetitive, monotonous assignments.

C. Teaching techniques or methods (positive relationships)

1. Assigning homework as well as seatwork.
2. Gaining the whole class' attention before beginning an explanation or giving directions.
3. Having and using science demonstrations or experiments.
4. Structured time in Math and Art relates to gains in reading and arithmetic computation and reasoning.

D. Teaching techniques or methods (negative relationships)

1. A high percentage of teacher structured time spent in reading groups.
2. When a child is stuck during reading, teacher asks another child to give him the word.

Discussion

In general, it appears that variables which measure control over the classroom are positively related to effectiveness. The teacher who is well-organized, who monitors the class regularly and nips potentially serious problems in the bud, and who has well established routines for handling everyday procedural matters tends to be more successful in producing learning gains. These data provide strong support for the observations of Kounin (1970).

It appears that a key factor in effective teaching is organizing the classroom environment so that there is maximum opportunity to learn. The daily routine and non-academic details are dealt with efficiently and kept to a minimum so that the task of learning can proceed, and this is especially important in the low SES schools. For giving directions or explanations there appears to be an optimum level of effectiveness. The teacher who dwells too long on details or becomes overly repetitive appears to have a detrimental effect, especially in the higher SES schools.

A few variables show significantly positive relationships in one group and significantly negative relationships in the other. For example, for Non-Title I children, delaying the child who doesn't understand until later when the teacher has time to devote to him shows a positive relationship with student gain, but in Title I this relationship is negative. Also when a child is stuck during a reading turn, for Title I, giving the child a clue unrelated to the meaning or sound of the word is negatively related but in Non-Title I the relationship is positive. A large amount of teacher structured time in reading groups appears as positive for Title I, but negative for Non-Title I.

Some cautions need to be stressed in interpreting these data, which are tentative and in need of replication for several reasons. First, data were based on only 2 to 4 observations per teacher. These frequencies are dangerously small given the probable day-to-day variation in teacher performance and the effects of situational factors such as finishing or beginning a unit, weather conditions, etc. Second, the probability data given to indicate the strength of relationships cannot be taken very seriously because the sample contained only 31 teachers and over 1000 measures were taken. This obviously violates assumptions underlying significance tests. Third, partly because of the low number of opportunities to observe the teachers, the observers' high-inference checklists and ratings show evidence of halo effect and logical error, so that some of the variables remain suspect in spite of high inter-observer agreement. Fourth, the Pearson r's reflect only linear relationships and do not take into account possible curvilinear relationships between predictors and criteria. We suspect that several variables will show such relationships. Fifth, the variances of the predictors should be examined, since unusually large or small variances will increase or minimize the chances for correlation. Sixth, several of the variables may be "proxy" variables; i.e., they may correlate with another variable which is the real predictor with the criterion in much the same way as income is correlated with education. We are aware that there may be other interpretations of the data depending upon one's own persuasion, hence the correlations themselves are presented in the tables with a minimum of interpretation.

Because of the limitations mentioned above, and because,even without these limitations, the data represent a unique set of findings which are not directly comparable to any other data, replication is being carried out with at least 15 observations per teacher and with some additional information gained in the form

of a personal interview with each teacher this year. This replication study will help determine which of the correlates of teacher effectiveness in producing learning gains are genuine and dependable. The correlates that replicate will then be experimentally manipulated to establish whether or not they are causally effective in producing student learning.

For other data from the first year's analyses, see Brophy and Evertson (1973) and Peck and Veldman (1973).

References

Brophy, J. Stability in teacher effectiveness. Report Series No. 77, Research and Development Center for Teacher Education, The University of Texas at Austin, 1972.

Brophy, J. and Evertson, C. Low-inference observational coding measures and teacher effectiveness. Paper presented to the Annual Meeting of the American Educational Research Association, New Orleans, La., 1973.

Brophy, J. and Good, T. The Brophy-Good dyadic interaction system, In A. Simon and E. Boyer (Eds.). Mirrors for Behavior: An Anthology of Observation Instruments Continued, 1970 Supplement, Volume A. Philadelphia: Research for Better Schools, Inc., 1970.

Emmer, E. Classroom observation scales. Research and Development Center for Teacher Education, The University of Texas at Austin, 1972.

Fuller, F. Concerns of teachers: A developmental conceptualization. American Educational Research Journal, 1969, 6, 207-226.

Kounin, J. Discipline and Group Management in Classrooms. Holt, Rinehart and Winston: New York, 1970.

Peck, R. and Veldman, D. Personal characteristics associated with effective teaching. Paper presented to the Annual Meeting of the American Educational Research Association, New Orleans, La., 1973.

Veldman, D. Comprehensive personal assessment system for teacher education programs. Research and Development Center for Teacher Education, The University of Texas at Austin, 1972.

Table 1. Correlations between Behavior Observation Scales[1] and Residual Gain Scores for Total Group, Title I and Non-Title Schools (decimal points omitted).

les	Word Knowledge (N = 31)		Word Discrimination (N = 31)		Reading (N = 31)		Arithmetic Computation (N = 31)		Arithmetic Reasoning (N = 22)		Rater[2] Agreement
	Total		Total		Total		Total		Total		
	TI	NT	TI	NT	TI	NT	TI	NT	TI	NT	
High Level of Student Attention	21		19		32*		25		15		88
	-02	23	17	10	20	27	17	30	33	14	
Teacher Often Addresses Questions or Problems to the Whole Class	22		36**		08		-09		-04		77
	01	21	29	31	-36	06	-19	-09	-50	-03	
Teacher is Task Oriented, Doesn't Waste Time	27		30*		31*		15		08		83
	28	21	47	12	17	24	16	14	23	06	
Frequent Pupil-to-Pupil Interaction (Class Relevant)	-02		-22		11		20		16		79
	-10	00	-35	-17	34	07	59**	05	17	16	
% of Time Teacher Lectures or Demonstrates	36**		40**		19		13		11		89
	36	33	22	41	-10	21	30	03	50	07	
Negative Affect: Criticism, Hostility	-12		-22		-18		-06		01		83
	11	-18	-22	-14	-09	-15	15	-21	14	01	
Positive Affect: Praise, Support	11		10		09		16		04		86
	-12	08	28	-15	-05	-06	-06	27	-56	06	
Requires High Level of Generalization, Inference, or Explanation	21		23		21		-07		-04		83
	23	15	39	04	-24	19	06	-12	-31	-05	

9. High Student Withdrawal, Passivity, or Aimless or Repetitive Behavior	-13		-11		-21		-10		-04		75
	25	-18	09	00	20	-19	-08	-10	-51	.05	
10. Clarity: Students Show Clear Understanding of Teacher Presentations	20		16		24		19		15		86
	-05	24	17	05	04	23	14	23	60*	11	
11. Enthusiasm: Teacher Shows Enthusiasm, Excitement, Enjoyment	-03		-04		05		01		-09		95
	-20	-13	15	-38	08	-20	-02	-01	-70	-10	
12. Convergent Questioning: Most Questions Have Clear-Cut Correct Answers	25		19		-02		-05		05		72
	11	28	18	18	16	-10	-16	-01	25	03	

[1] 5-Point scales rated intermittently during each visit and averaged across observers.

[2] % agreement within 1 point by two independent raters.

* $p < .10$

** $p < .05$

*** $p < .01$

Table 2. Correlations between Teacher Behavior Checklist Variables[1] and Residual Gain Scores for Total Group, Title I and Non-Title Schools (decimal points omitted).

Checklist Variables	Word Knowledge (N=27-30) Total		Word Discrimination (N=27-30) Total		Reading (N=27-30) Total		Arithmetic Computation (N=27-30) Total		Arithmetic Reasoning (N=19-21) Total		Rater[2] Agreement		
	TI	NT	TI	NT	TI	NT	TI	NT	TI	NT	Both	One	None
A. Methods of Handling Catch-Up Work													
1. No Remediation;													
Child Skips Missed													
Work	-07	-09	-09	-03	04	26	11	20	-02	-11	0	2	108
2. Child Must Make Up													
Work but Is Not													
Given Help	-03	25	07	07	-07	07	-01	15	-13	08	1	5	104
3. Teacher Explains													
Work and Has Child													
Do Part of It.	26	16 09	18	24 -07	17	24 03	19	33 -02	-07	-21	1	13	96
4. Another Child Is													
Assigned to Help	11	13	04	13	15	31	14	35	-19		0	6	104
5. Child Put in Slower													
Group Temporarily	-13	-27	-21	-33	-13	-18	18	31	-04	-21	1	0	109
6. Other	-03	00	-17	-25	-02	07	21	37	07	33	0	1	109
B. Rules Regarding Physical Movement													
1. Must Always Get													
Permission to													
Leave Seat	06	33 -01	16	16 29	05	17	09	19	09	14 12	2	10	97
2. One at a Time													
Without Per-													
mission	35*	36	60***	69***	31*	32	25	31	27	28	0	1	108

3. As Many as 4 or 5 Without Permission	04	04	-08	-33*	-10	0	5	104
	-10 30	04 26	-22 20	-61** -	-76 08			
4. Can Go Quietly to Specified Places without Permission at Any Time	02	-16	06	05	-03	23	20	66
	-50* 15	-29 -29	-20 06	06 12	47 -			
5. No Restrictions	06	05	-03	03	01	4	12	93
	42 -03	29 06	29 -10	34 -28	70 -14			
6. Some Children Allowed Free Movement but not Others	-25	-18	-22	07	-03	1	1	107
	-27 -26	-33 -10	-18 -25	31 -11	-21 -			
7. Only Monitors Allowed Free Movement	07	-03	13	15	--	0	1	108
	26 -	04 -	43 -	26 -	- -			
8. Other	-39**	-34*	-27	-27	-37*	1	3	105
	- 43 -49	-45 -27	-23 -34	-23 -35	- -39			
C. Punishments Used by Teacher								
1. Stay after School	07	07	30*	16	13	5	8	96
	28 -	22 01	78** 11	33 01	16 16			
2. Spanking	19	11	-07	-05	02	1	3	105
	54 09	30 10	-09 09	07 01	14 -			
3. Writing Sentences on Board	--	--	--	--	--	0	0	109
4. Isolation within the classroom	10	10	12	01	-04	4	14	91
	27 03	09 09	37 -02	01 01	06 -05			
5. Removal from the Classroom	-06	-11	-06	-09	-02	9	7	93
	21 -16	-04 -13	24 -17	-05 -11	21 -07			
6. Note to Parents	10	12	11	01	--	0	1	108
	- 09	- 10	- 09	- 01	- -			
7. Send to Principal	-05	00	-11	00	01	3	4	102
	18 -11	-07 -02	-30 -10	19 -08	-07 03			

											19		
8. Extra Reading, Math, Etc. Work	-20		-03		-09		15		11		0	5	104
	-27	-23	-33	-03	-18	-14	31	11	-21	13			
9. Peer Pressure (e.g. "You lost the race for your group.")	08		-05		09		03		07		1	15	93
	33	-03	-05	-16	-25	08	24	02	18	01			
10. Scolding	-09		00		03		13		16		12	23	74
	02	-11	-13	24	30	01	37	-07	57	13			
11. Discussion of Incident (No Scolding)	16		-05		22		02		21		8	14	87
	-25	38	-08	-04	-	37	-30	31	-80	44*			
12. Other	23		08		32*		10		09		1	13	95
	21	20	08	01	22	29	05	12	-41	13			
D. Rewards Used by Teacher													
1. Classmates Clap or Cheer	-06		08		-02		-23		-29		3	16	90
	-41	03	-18	16	-24	02	-31	-19	-81**	-24			
2. Special Privileges	-08		-12		16"		06		-11		0	7	102
	26	-27	04	-31	43	-02	26	-09	-	-13			
3. Waiver or Reduction of Assignments	-35*		-19		-15		-29		-51**		0	1	108
	-	-41	-	-27	-	-21	-	-37	-	-53**			
4. Symbols (Stars, Smiling Faces, etc.)	00		-02		-13		03		-16		3	13	93
	21	-08	30	-19	10	-24	31	-14	27	-21			
5. Token Redeemable for other Rewards	00		-10		-01		12		13		2	1	106
	-	-03	-	-19	-	-08	-	15	-	14			
6. Concrete (Candy, Money, Prizes)	-06		-19		-19		13		10		0	3	106
	-40	16	-59*	17	-35	-06	08	20	-21	19			
7. Jobs (Monitor, Helper, Clean Erasers)	07		08		14		00		09		2	5	102
	03	10	-06	22	22	14	-06	05	-85*	21			
8. Public Recognition (Gets to Read or Work Problem on [illegible]d)	17		11		26		-11		07		8	16	85
	09	18	-08	15	-26	40	-39	03	-26	10			

9. Other	-13		-15			-24		-10			05			0	8	101
	-36	16	-22	17		-31	-06	-29	20		-30	19				
E. Appropriateness of Assignments																
1. Too Short or Easy	-23		-16			-06		-03			-09			13	18	77
	-36	-16	-23	-02		18	-11	27	-35		54	-31				
2. Boring, Repititive, Monotonous	-24		-04			-20		-13			-15			7	20	81
	27	-45*	30	-05		29	-33	08	-33		02	-23				
3. Too Hard: Students Can't Get Started or Continually Need Help	-11		-15			-21		00			-02			3	17	88
	03	-13	-03	-07		-06	-19	02	-		-71	10				
4. Continues Activity Too Long, until They Get Boring	-27		-22			-24		-10			-05			5	17	86
	-15	-30	-25	-12		-09	-23	16	-26		-20	-04				
5. No Inappropriate Assignments	30		22			40**		34*			25			11	17	80
	04	35	08	13		20	40	28	44*		-41	33				
F. Distractions: What Do Students Do When not Working?																
1. Use Washroom	06		-10			-01		-02			-23			14	18	76
	39	-03	17	-19		22	-04	32	-29		-17	-27				
2. Repeatedly Get Supplies for Free Time Activities	-05		-16			-05		03			13			6	10	92
	-35	08	-59**	11		-21	07	08	-01		05	17				
3. Watch Reading Group or other Activity	-35**		01			-34*		-41**			-33			3	20	85
	-42	-33	-18	12		-27	-35	-31	-46*		-21	-34				
4. Talk	10		-08			07		-13			-04			33	20	55
5. Play	00		-23			-02		16			15			12	20	76
	01	09	-36	-03		21	-	39	01		82**	13				
6. Daydream	-16		-25			-24		-15			03			11	19	78
	-24	-09	-34	-11		-49	-04	-48*	11		-58	16				
7. Ask for Help or Look More Closely at Work on Board	03		-07			17		22			25			6	18	84
	-	06	-12	-01		30	17	45	09		34	26				

3. Disrupt other Students	-03	08	-08	-09	-12	-08	-08	-01	-13	00	-03	02	-04	-17	-03	15	19	74
9. Other	-19	26	-38	-17	-19	-07	-36**	-23	-40	-19	29	-59***	-20	18	-32	0	7	101
Student Attitudes toward the Teacher																		
1. When Having Trouble Students Concentrate or Seek Help	-13	-11	-19	-09	06	-27	02	25	-15	-08	13	-24	-17	-12	-20	13	30	62
2. When Having Trouble, Students Merely Copy from Neighbor	-28	-68***	-10	-37*	-68***	-10	-21	-14	-14	-10	-16	-06	-06	-43		8	16	81
3. Students Work as Well When not Watched as When Watched	09	-26	17	12	-02	09	25	-05	31	13	-07	25	23	-59	35	11	16	78
4. Students "Act Up" When Unwatched	-06	04	-05	-19	-36	-01	-05	22	-07	07	22	-03	07	48	05	11	18	76
5. Students Seem Amused by Teacher	-06	16	-13	04	15	02	02	30	-07	10	20	04	06	02	10	4	8	93
6. Students Seem to Fear Teacher	10	29	08	21	19	43*	22	67***	05	21	31	14	21	48	21	0	5	100
7. Students Seem to Respect Teacher	13	23	02	16	43	-17	01	-29	-01	-08	-30	07	-04	-62	04	28	12	55
Free Time Materials Available (Not Necessarily Used)																		
1. Books	-25	-10	-45*	-09	-01	-25	-34*	-28	-54**	-32*	-38	-30	-39*	-33	-46*	41	16	45
2. Learning Centers (Any)	11	-03	06	19	08	09	19	17	06	-05	-38	09	07	-22	12	9	21	72

3. Listening Centers	-02			-01			00			-21			-28			19	14	69
	12	-09		18	-13		28	-14		-29	-16		-01	-33				
4. Visual (Picture Files, Filmstrips)	07			15			02			-08			-10			4	19	79
	74**	-12		43	-04		18	-14		31	-25		56	-16				
5. Science Demonstrations or Experiments	15			01			10			-04			00			5	16	81
	08	14		-20	02		15	-02		07	-14		-	-.01				
6. Other Learning Centers	18			-04			-02			12			21			1	9	92
	16	18		-13	-02		-25	04		-02	19		28	21				
7. Coloring Pictures	03			-23			11			08			01			7	27	68
	-26	15		-30	-19		01	17		08	07		-13	04				
8. Painting, Art Activities	-20			-29			-11			-09			-29			8	19	75
	-12	-25		-22	-35		33	-33		.21	-34		-07	-35				
9. Games (Any)	-03			-04			-02			-06			-08			14	22	66
	-43	05		-30	-03		-19	-05		-27	04		-09	-08				
10. Instructional Games	-16			-24			-10			-10			-08			23	18	61
	-56**	-06		-45	-21		-09	-18		-45	14		-54	-02				
11. Non-instructional Games	-13			-21			-15			-06			--			14	24	64
	-51*	-01		-38	-14		-13	-16		-37	13		-21	03				
12. Aquarium, other Looking Exhibits	02			03			-15			-20			00			18	16	68
	39	-11		34	-13		-10	-17		-28	-14		-03	01				
I. Free Time Materials Observed in Use																		
1. Books	10			19			-08			-27			-20			25	16	61
	12	06		34	04		-01	-21		-37	-23		-28	-21				
2. Learning Centers (Any)	20			11			28			02			10			4	15	83
	-06	20		16	-03		31	19		-34	12		-36	14				
3. Listening Centers	09			-03			19			-07			-28			9	17	76
	04	08		16	-16		60**	02		-	-11		-04	-31				
4. Visual (Picture Files, Filmstrips)	40**			28			26			16			13			3	14	85
	58***	33		31	20		16	23		22	13		38	11				
5. Science Demonstrations or Experiments	31*			20			29			09			16			0	15	87
	-09	35		-03	14		-35	32		-25	21		-	18				

6. Other Learning Centers	19		-05		01		12		24		1	10	91
	16	20	-13	-04	-25	08	-02	20	28	24			
7. Coloring Pictures	13		-11		09		00		03		5	25	74
	-24	30	-26	-02	-13	21	-12	08	-47	13			
8. Painting, Art Activities	-11		-16		01		-05		-27		5	15	82
	-02	-18	-04	-27	37	-19	29	-32	-07	-36			
9. Games (Any)	-02		-14		01		-06		-08		6	23	73
	-44	11	-26	-13	-11	02	-21	03	-12	-08			
10. Instructional Games	-01		-12		13		03		-10		12	16	74
	-32	08	-06	-19	45	-02	-08	10	-54	-04			
11. Non-Instructional Games	-06		-16		-03		-05		-06		12	19	71
	-38	07	-26	-11	03	-05	-23	10	-06	-07			
12. Aquariums, Other Looking Exhibits	11		12		12		-14		-07		4	13	85
	42	02	70***	-10	26	-01	-11	19	06	-08			
J. Use of Peer Tutoring	-11		-12		-02		-09		-14		7	15	87
K. Assigns Homework besides Seatwork	23		13		34*		19		08		4	23	82
	-26	38	-29	25	-04	46*	19	19	-32	13			
L. Teacher Sometimes Underreacts to Control Problems, so Serious Problems Go Unresolved	-13		-07		-11		03		05		6	13	90
	-35	-06	-52*	21	-12	-12	17	-09	05	06			

---

[1] Each of two observers completed the checklists twice; scores obtained by summing.

[2] These are the raw rater agreement data: "Both" means that both observers checked the item; "One" means that one checked the item and one did not; "None" means that neither checked the item. Minor differences in the totals occur because observers occasionally felt completely unable to rate on a given item.

* $p < .10$

** $p < .05$

*** $p < .01$

Table 3. Correlations between High-Inference Ratings[1] and Residual Gain Scores for Total Group, Title I and Non-Title Schools (decimal points omitted).

Ratings Variables	Word Knowledge (N=24-28)		Word Discrimination (N=24-28)		Reading (N=24-28)		Arithmetic Computation (N=24-28)		Arithmetic Reasoning (N=18-20)		Rater[2] Agreement
	Total		Total		Total		Total		Total		
	TI	NT	TI	NT	TI	NT	TI	NT	TI	NT	
1. Typical Affectionateness Level	26		14		14		21		11		86
	-21	42	-01	19	-25	26	-03	35	-49	17	
2a. Most Intense Affection Expression Observed	14		01		22		28		14		90
	-52	44	-40	16	-04	24	20	30	-53	22	
2b. Most Intense Negative Affect Observed	-17		-18		-06		-06		-15		75
	11	18	37	00	-13	08	-28	42	-61	24	
3. Solidarity with Class: Teacher Identifies, Promotes "We" Feeling	22		16		16		01		06		84
	-25	39	15	08	-05	14	-34	28	-83**	20	
4. Patient and Supportive When Correcting	20		04		10		-03		03		86
	-14	32	11	-15	-25	14	-28	15	-51	-08	
5. Students Allowed Choice in Assignments	-01		00		-06		-20		-24		82
	16	-14	31	-22	33	-33	-22	-26	21	-37	
6. Accepts Student Ideas and/or Integrates them into Discussion	04		06		09		-13		-21		79
	-16	08	18	-13	06	-04	-20	-16	-84*	-28	

7. Admits Own Mistakes; Laughs at Self or Uses Occasion to Teach or Motivate	04		12		-09		-26		-19		84
	44	-17	78***-30		-02	-23	-34	-22	34	-30	
8. Usually Bends Close, Gets Down to Child's Level	38**		28		20		-04		14		71
	.25	47*	46	11	-10	30	-35	24	-31	24	
9. Goes to Seats to Check Work; Doesn't Stay at Desk	-03		-12		22		-13		-14		83
	-33	01	-12	-20	21	16	-06	-25	-53	-14	
10. Usually Speaks to Individuals rather than Whole Class	-13		-04		06		14		02		84
	-13	-22	-20	-01	11	-04	35	-12	-26	02	
11. Uses Advance Organizers in Introducing Activities	30		19		11		03		02		71
	-01	39	14	15	-37	22	-23	18	-73	13	
12. Gives Complete, Detailed Instructions; Prevents Errors before they Happen	32*		20		12		-17		04		67
	01	36	33	02	-44	18	-72***13		-59	10	
13. Students Eager to Respond; No Fear	17		24		27		19		02		86
	27	03	32	11	14	20	15	15	07	-02	
14. Teacher Waits Patiently if Student Doesn't Respond Promptly	13		02		-01		-19		-13		73
	21	06	44	-31	-25	01	-34	-03	-67	-13	

15. Non-Competitive Atmosphere; No Signs of Eagerness to See Others Fail	-12		-20		-19		-06		-17		50
	13	-13	01	-24	03	-10	16	-22	11	-27	
16. Students Allowed to Work In Cooperative Groups	10		-07		15		00		17		71
	-34	26	-20	00	27	14	-12	10	-47	24	
17. Teacher Recognizes Good Thinking Even When It Doesn't Lead to "Right" Answers	11		-07		03		-19		-22		83
	14	02	29	-41	-04	-11	-26	-23	-46	-26	
18. Democratic Leadership Style: Students Share In Planning and Decision Making	20		09		25		-07		-17		81
	02	16	26	-16	07	14	-03	-23	-63	-25	
19. Few Restrictions on Students During Seatwork Periods	10		-06		08		08		01		69
	-02	34	02	00	19	21	28	-05	30	-02	
20. Students Expected to Care for Needs Without Getting Permission	38**		15		29		17		26		66
	22	41	29	-01	26	21	13	15	57	22	
21. Teacher Concerned with Substantive Content, not Form, of Student Responses	08		-06		-10		23		14		75
	51	-10	33	-31	08	-15	47	-01	62	05	

22. Teacher Stresses Factual Realism, Rejects or Corrects Childish Idealism	06		-07		-10		-17		-16		86
	26	-04	40	-28	00	-20	-40	-04	-36	-19	
23. Teacher Credibility: Students Seem to Believe and Respect Teacher	20		11		16		09		04		76
	-08	27	17	-01	-16	19	-17	27	-61	12	
24. Showmanship: Teacher Is Melodramatic, Expressive, Gushy, Emotive	-11		09		-07		24		04		80
	-18	-15	15	15	30	-30	58**	-02	19	02	
25. Teacher Gets Attention before Starting, Doesn't Try to Talk over Din.	30		27		32*		-01		00		82
	17	33	41	10	-05	45*	-30	20	-66	07	
26. Chaotic, Unplanned, Poorly Scheduled	-25		-38**		-20		05		-24		78
	-33	-22	-48	-33	12	-42	46	-48*	20	-49*	
27. Teacher Seems Confident, Self-Assured	22		11		02		06		-04		90
	42	14	44	-07	-12	07	05	06	03	-07	
28. Politeness: Teacher Regularly Says "Please," "Thank You," etc.	02		00		02		00		-02		98
	02	-03	29	-26	11	-12	-15	12	-43	02	
29. High Concern about Achievement	23		15		09		-05		-07		88
	16	19	31	-06	-24	07	-31	08	-61	00	
30. Room Is Attractive	44**		35*		32*		04		11		74
	36	45*	55*	14	34	17	-10	05	08	10	

31. Teacher Gives Much Encouragement to Students	05		05		11		-07		-08		88
	-16	04	24	-28	06	-09	-26	03	-52	-08	
32. Room is Uncrowded	56***		40**		53***		46***		43**		74
	43	59**	21	41	03	62***	40	50**	60	48*	
33. Teacher explains Rules or Decisions When Reasons Aren't Obvious	40**		30		19		02		06		87
	32	41	53*	07	-08	18	-25	16	-64	10	
34. Teacher Well Organized, Prepared	46***		32*		21		-03		14		84
	50	42	55*	09	-21	25	-47	27	-36	25	
35. Teacher Regularly Monitors Class, Knows What's Going on	35*		32*		23		00		00		86
	34	31	53*	08	-04	20	-25	14	-40	04	
36. Smooth, Efficient Transitions, Little Time Wasted	49***		36**		35*		13		14		70
	50	47*	64**	15	08	37	-08	25	-24	19	
37. Monitors Determined "Automatically" by a Systematic Procedure	37*		37*		50***		27		18		74
	41	29	56**	16	41	44*	12	33	36	16	
38. "Busy," Cluttered Classroom	-02		05		12		-20		-09		86
	-26	-	05	-07	22	-13	-16	-43*	10	-25	
39. Students Compliant, Obedient	27		23		23		12		09		94
	-04	36	24	16	-19	33	-21	34	-84**	20	
40. Teacher Gives Overly Explicit, Repetitive Directions	-12		00		-23		-53***		-37*		74
	-07	-21	13	-13	03	-51**	-62**	-17*	-50	-45	

41. Well Established Routines Minimize Interruptions; Room Runs "Automatically"

29		22		27		06		08	82
-01	34	32	07	-01	26	-27	25	-83**17	

---

[1] First three scales have 13 points; all others have 5. Scores obtained by summing across two raters.

[2] % Agreement across rater pairs; within two points on first three scales, within one point on all others.

* $p < .10$

** $p < .05$

*** $p < .01$

Table 4. Correlations between Percentage Estimate Variables[1] and Residual Gain Scores for Total Group, Title I and Non-Title Schools (decimal points omitted).

Percentage Estimate Variables	Word Knowledge (N=22-31) Total TI NT	Word Discrimination (N=22-31) Total TI NT	Reading (N=22-31) Total TI NT	Arith. Computation (N=22-31) Total TI NT	Arith. Reasoning (N=15-22) Total TI NT	Rater[2] Agreement Both One None
A. Time Utilization[3]						
1. % Total Time Structured by Teacher	-09 11 -19	09 18 03	-30* -10 -43*	-21 -09 -30	-16 -02 -20	-- -- --
2. %Structured Time in Language Arts	-04 -50* 14	-08 -34 07	02 -26 17	03 -15 15	17 -06 21	-- -- --
3. % Structured Time in Math	24 13 25	05 -10 -08	36** 21 34	32* 25 44*	25 36 29	-- -- --
4. % Structured Time in Art	30* -04 39	12 -25 24	43** -09 58**	25 07 34	45** 42 46*	-- -- --
5. % Structured Time in Spelling	-01 09 -09	10 -08 15	-10 -21 -11	-13 17 -36	-21 01 -30	-- -- --
6. % Structured Time in Reading Groups	-16 38 -33	-09 54* -34	-22 39 -40*	-05 52*-42*	-33 39 -43*	-- -- --
7. % Structured Time in Social Studies	-05 -05 -10	09 10 01	04 26 -09	01 -43 14	01 -59 05	-- -- --
8. % Structured Time in Transitions	09 -02 16	03 -08 13	-05 -24 06	-01 -15 11	12 29 14	-- -- --
9. % Structured Time in Morning Routine	-09 02 -13	-14 -26 -08	-03 13 -09	02 32 -19	-17 28 -27	-- -- --
10. % Structured Time in Special Activities	-09 -03 -08	-03 02 04	-25 -19 -23	-29 -59**02	-12 -48 -06	-- -- --
B. Methods Used to Call for Attention						
1. Says Nothing, Waits for Quiet	19 -05 27	29 22 26	25 -27 45*	-09 -45 17	-03 -77* 07	21 6 15
2. Raps Desk Lightly, Uses Normal Voice	05 -35 31	08 -07 26	15 27 13	20 16 26	22 -05 25	1 6 35
3. Gimmick (light flick, bell, clicker)	-25 09 -36	-10 -14 -22	-29 06 -46	-20 34 -34	-17 33 -19	5 8 39
4. Raises Voice Over the Din	-14 -25 -06	-26 -32 -14	-18 -17 -11	08 07 12	14 68 10	22 32 8

5. Raises Voice and Singles Out Individuals	-06 -03 -07	-05 -07 -03	-06 -12 -02	-14 -17 -12	-07 38 -12	12	23	17
6. Shouts, Becomes Angry, or Scolds Class	03 42 -15	-09 06 -02	02 45 -15	17 39 -09	04 32 --	5	5	42
7. Shouts, Becomes Angry, or Scolds Individuals	-07 40 -30	-12 04 -13	-12 23 -27	10 36 -16	-02 20 -07	5	6	41
8. Whispers or Speaks Softly to Nearby Pupils (at first)	15 31 15	23 38 27	10 29 06	21 27 19	12 -- 12	0	5	47
9. Other (includes any method not listed above)	09 -03 12	-12 -19 -10	05 41 -05	17 42 07	10 57 07	3	12	37
C. Estimated % of Students Paying Attention[4]	18 13 16	23 47*-01	23 08 21	03 -23 21	07 -25 15	50	0	0
D. What Does the Teacher Do When a Child Doesn't Understand?								
1. Stops What She's Doing, Explains	-16 -32 -16	-02 01 -11	-24 -55**-22	-45** -58* -41	-45** -43 -49*	29	11	1
2. Delays Child then Explains Later	29 -66** 52**	03 -44 17	20 -41 36	11 -42 35	39* -53 46*	21	16	4
3. Delays, but then Fails to Follow Up	-23 -30 -20	-13 -23 -07	-15 03 -20	05 33 -11	03 20 02	3	10	28
4. Asks Another Child to Explain	17 58**-02	41** 59** 27	21 17 18	03 06 01	11 83** 03	8	22	10
5. Scolds Child for Not Understanding	-01 26 -17	-13 -05 -02	06 47*-08	16 33 -03	03 03 07	3	10	27
6. Encourages Child but Give No Help	-20 -06 -31	-15 -08 -22	08 34 -08	22 31 16	-05 -10 -04	3	14	23
7. Refuses Help ("You're on your own.")	08 51*-13	-03 10 -04	04 35 -08	22 26 25	16 21 19	2	13	25
8. Sends Child to Aide or other Adult	-01 09 -07	08 28 -01	-04 -06 20	07 08 17	15 55 11	-	1	39
9. Other (includes any method not listed above.)	-09 -12 -08	-34* -38 -38	-13 -02 -23	25 48*-02	-02 23 -18	0	4	36
E. Teacher Goes to Child's Desk to Give Help, Doesn't Stay at Desk	19 45 09	28 39 20	08 26 -03	22 53* 02	15 67 09	50	0	0

F. What Teacher Does When Child is Stuck While Reading in Reading Group

1. Gives Word	11	01	13	34*	35	26	3	0
	12 17	-10 16	-08 31	29 40	79* 34			
2. Gives First Sound or Syllable	10	01	03	08	-08	12	9	8
	11 06	36 -21	17 -09	-06 14	-67 -03			
3. Child Starts Sentence or Paragraph Over	19	-07	-09	01	09	0	12	17
	18 22	-10 -05	-37 02	-13 09	19 09			
4. Gives Context Clue or Definition	-24	-09	-23	-39**	-29	4	5	20
	-30 -23	-20 -04	-12 -29	-17 -52**	-94***-22			
5. Asks Another Child to Give Word	-32*	04	-26	-48**	-51**	15	7	7
	-14 -44*	-07 00	-13 -40	-29 -61***	-02 -56**			
6. Gives Clue Unrelated to Sound or Meaning ("It's one of our new words.")	30	04	37*	09	26	1	7	21
	-40 45*	-36 18	-10 46*	-21 21	-95*** 35			
7. Tells Child to Skip, Go to Next Word	-26	-17	-10	13	02	0	2	27
	-22 -30	-27 -13	-03 -13	34 00	-20 06			
8. Other (includes any method not listed above)	03	08	18	-03	01	0	2	27
	18 --	28 --	63** --	-02 --	06 --			

[1]Two observers estimated percentages for each appropriate category; scores obtained by averaging.

[2]These are the raw rater agreement data: "Both" means that both observers entered a % > 0 on the item; "One" means that one entered a % > 0 and the other entered "0;" "None" means that both observers entered "0." Differences in totals occur because observers were not always able to estimate with any confidence.

[3]Time utilization data were computed from information on the classroom coding sheets concerning the starting and ending of activities. Reliability coefficients were not computed here because agreement was near-perfect; minor differences were handled by averaging across observers.

[4]Pearson r was computed for this variable, since both observers made % estimates in every case; $r = .67$.

[5]$r = .50$

* $p < .10$

** $p < .05$

*** $p < .01$