ABSTRACT
                This report presents a portion of the evaluation of
the planned variation of Project Follow Through, in which increased
understanding of education which is functional for disadvantaged
children has been a major concern. This segment of the evaluation has
the following two objectives: (1) to describe in behavioral terms the
differences among programs as observed in classrooms, and (2) to
relate these dimensions to pupil growth. Observation was carried out
in classrooms from each of seven programs and a comparison sample.
Observers recorded such things as the teacher's classroom management
techniques, expression of affect, the social-emotional nature, and
the cognitive level of the interaction. Results were then factor
analyzed, reduced to factor scores, and the factor scores tested for
differences between programs by the multiple range test. The factor
scores were also related to classroom mean pupil regressed gain which
had been factored to identify different levels of complexity.
Although there were inconsistencies, the linear relations between the
classroom measures and pupil gain suggested that some factors, such
as effective personal communication and intrinsic motivation, related
to all levels of complexity, while others were specific to a certain
level. In general, there seemed to be a trend that abstract growth
was related to teaching that was less controlled and less focused but
had some structure, while skill growth was related to more focus and
structure, with concrete growth positively related to still more
highly focused teacher behaviors and negatively related to extreme
pupil freedom. (Author/RC)

FOLLOW THROUGH CLASSROOM PROCESS MEASUREMENT*

By

Robert S. Soar

INSTITUTE FOR DEVELOPMENT OF HUMAN RESOURCES

COLLEGE OF EDUCATION
UNIVERSITY OF FLORIDA
GAINESVILLE, FLORIDA 32611

June, 1971

2

Acknowledgments

In any project, the help of many people is required. But when that help is given enthusiastically and unstintingly, special thanks are due. That was the case in this project.

Ruth Soar has been codirector in every way but in name, collaborating in application and report writing, wrestling with budgets, collecting data, planning and programming data analyses, and interpreting the results. Marjorie Ragosta has been central in the work of the project since its beginning. She has served as observer, coder and observer trainer and supervisor, data processor, field supervisor and project manager. She has organized and coordinated the work of field staff and coders, and coordinated the project with school systems. Mary Webb has been observer and provided valuable "back up" as trainer, supervisor and "house mother" to the staff of student coders who worked diligently to decipher tapes which were sometimes difficult.

Other observers who have often worked days and traveled nights were: Bill Blair, Pat Quigley, Sister Maura Smith, Kathy Gussow, Judy Godfrey, and Linda Stebbins.

The data processing was coordinated by Linda Stebbins, supported by Keith Brown, Earl Blekking, and Bill Graves. Jan Larson, Barbie Hollien, and Denise Stevens have dealt with the myriad administrative and clerical details that accompanied the project.

Finally, Follow Through coordinators and school system supervisors have smoothed the way, planning schedules and arranging with teacher for our visits. They and their principals have welcomed us and made us feel at home. But the real heroes in the front lines have been the teachers, who despite the many demands on their time and energy, have welcomed us into their classrooms, and

given us the freedom to gather the most meaningful picture of the classroom we had the skill to record.

To all, a heartfelt thanks!

Contents

Contents (Continued)

List of Tables

List of Tables

List of Figures

10

Introduction

As part of the evaluation of project Follow Through, the Institute for Development of Human Resources of the University of Florida (IDHR) assumed responsibility for collecting observational data in a sample of classrooms representing a number of experimental programs. Three waves of data have been collected: 70 classrooms each in the winters of 1969 and 1970; and 289 in the winter of 1971. The results for the first two years are reported here.

Two sets of goals lay behind this effort:

1. To describe in behavioral terms the differences among the programs as observed in the classrooms, and

2. To relate these behavioral dimensions to pupil growth.

As work advanced toward these goals, several subgoals emerged and will be described later.

The observational measures were not focused directly on the identification of sponsor objectives and the development of items to represent them. Rather, they were selected from already existing instruments and represented a broad conception of classroom interaction as it has been developed over the past years. The instruments ranged from one with very extensive research background to two with some previous use, to one which was newly developed from work of others.

To enable study of relations between measures of classroom observation and pupil growth, Stanford Research Institute (SRI), the principal outside evaluator of Follow Through procedure, provided test data on pupils.

Procedure

Sample

Seven programs were selected in which at least eight classrooms could be observed which seemed to represent the diversity of programs present in Follow Through (although the latter criterion was a subjective and uncertain one).

Two comparison classrooms were selected from the same settings in which the programs were located, in the hope of equating, in a rough way, system-related variance for program and comparison classrooms. Insofar as possible, settings and programs were selected for observation where SRI was collecting complete data from pupils.

Each year a total of 70 classrooms was observed: eight program and two comparison classrooms from each of seven programs. The second year the attempt was made to observe the same teachers as far as possible. Exceptions to this were numerous, however, as a consequence of teacher resignations and promotions and, in one case, the need to replace an entire community which had been sampled the first year. The programs and the locations for the samples for the two years are shown in Tables 1 and 2. The way in which the sample was selected sets two limitations on the meaningfulness of the results. One limitation is that SRI collected data in classrooms suggested by program sponsors as ones in which the program was being well implemented. This designation was made by sponsors at the beginning of the first year and on the basis of little information, and in some cases the selection turned out later not to have been the best choice. The other limitation, a function both of sample and of schedule, was that the programs had been in operation in some of these locations for only about three months at the time the first year's observation began. As a consequence, some classrooms did not represent their programs well, and some programs had had relatively little time to become well implemented at the time of data collection.

On the other hand, this situation was not as atypical as it may seem. Experience the second year indicated that it was not unusual for half of the teachers in a program to be replaced because of resignations and promotions.

Table 1

Sample by Program and Community - 1969

Community	Bank Street	Becker Engel-mann	Educ. Devel. Center	Gotkin	Nim-nicht	Parent Educ.	Tucson
Atlanta, Ga.				4+2*			
Berkeley, Calif.					3+1		
Burlington, Vt.			2				
Cleveland, Ohio					2+1		
Duval Co. (Jax), Fla.						2+1	
Duluth, Minn.					3		
East St. Louis, Ill.		3+1					
Ft. Worth, Texas							3
Jonesboro, Ark.						3	
LaFayette, Ga.							3+1
Laurel, Delaware			2+1				
New York City				4			
Philadelphia, Pa.	.3+1		4+1			5+1	
Racine, Wis.		4+1					
Rochester, N. Y.	2+1						
Tuskegee, Ala.	3						
Vincennes, Ind.							2+1
Washington, D. C.		1					
Program Total	8	8	8	8	8	8	8
Comparison Total	2	2	2	2	2	2	2

*The second entry represents number of comparison classrooms.

Table 2

Sample by Program and Community - 1970

Community	Bank Street	Becker Engelmann	Educ. Devel. Center	Gotkin	Nimnicht	Parent Educ.	Tuscon
Atlanta, Ga.				4+2*	3+		
Berkeley, Calif.					3+1		
Burlington, Vt.			2+1				
Duval Co. (Jax), Fla.						2+1	
Duluth, Minn.					3		
East St. Louis, Ill.		3+1					
Ft. Worth, Texas							3
Jonesboro, Ark.						3	
LaFayette, Ga.							3-1
Laurel, Delaware			2				
Lebanon, N. H.					2+1		
New York City				4			
Philadelphia, Pa.	3		4+1			3+1	
Racine, Wis.		3+1					
Rochester, N. Y.	2+1						
Tuskegee, Ala.	3+1						
Uvalde, Texas		2					
Vincennes, Ind.							2+1
Program Total	8+2	8+2	8+2	8+2	8+2	8+2	8+2
Teachers observed both years	1+0	3+0	8+1	6+2	1+0	8+0	0+1

*The second entry represents number of comparison classrooms.

Since many programs were approximately doubling each year at that point, three-quarters of the teachers in a program were new in many cases.

Classroom Observation Measures

Florida Affective Categories (FLAC) and Florida Climate and Control System (FLACCS) - FLAC (Figure 1), which was used the first of the two years, was a modification of the South Carolina Observation Records, which was developed to be used as a parallel observation schedule with Interaction Analysis (Soar, 1966). The original instrument drew heavily on the Hostility-Affection Schedule (Fowler, 1962), and the earlier versions of the Observation Schedule and Record (Medley and Mitzel, 1958, private communication). The rationale of the instrument, overall, was the development of a schedule which would emphasize behavior which Interaction Analysis did not record. Among these were the nonverbal expression of affect in the classroom, physical movement of teacher and pupils, the groupings found in the classroom, and the extent to which individuals or groups were central in classroom activities. The nonverbal affect expression in the classroom seemed important to observe for two reasons: it seemed reasonable to assume that the teacher might be more successful in monitoring her verbal behavior than her nonverbal behavior when an observer was present (some evidence in earlier studies indicated this); and that relations between pupils might be an indicator of the emotional and social climate of the classroom which would be less likely to change with the presence of an observer than teacher behavior.

The Hostility-Affection Schedule allotted separate sections to the eight combinations of behavior: teacher-pupil, verbal-nonverbal, and supportive-non-supportive. If, for example, the teacher raps on her desk for order, pushes a child who is out of line, or waits impatiently for whispering to stop before proceeding, it would be tallied as teacher, nonverbal, nonsupportive. On the

Figure 1 - Florida Affective Categories (FLAC), 1969
(Class Form)

Teacher _____ Grade _____ Observation _____
Number

School _____ City _____

Date _____ Hour _____

Children's Original Art Work

Abundance of varied and vivid art-work	Quite a few varied and vivid art products	Some children's varied and vivid art products	A few children's original varied and vivid art products	No children's original art products
5	4	3	2	1

Room Displays and Artifacts

Most are clearly related to children's subcultural background	Quite a few are related to children's subcultural background	Some are related to children's subcultural background	A few are related to children's subcultural background	None show any relationship to children's subcultural background
5	4	3	2	1

16

Figure 1 - Continued

Tot	1	2	3	Teacher
				Leads singing, exercise, games
				Leaves, enters room
				Moves freely
				Withdraws from class
				Supervises pupil closely
				Immobilizes pupils
				Pnshs. class for ind. pupil
				Ignores pupil
				Refuses to attend pupil
				Attends pupil briefly
				Attends pupils in succession
				Attends pupil closely
				Attends simult. activities
				Meets pupil's needs
				Directs with song or music

Pupil	1	2	3	Tot
Seeks information				
Gives information				
Follows cognitive plan, T. pblm.				
Follows cognitive plan, P. pblm.				
Absorbed in work				
Intermittent work				
Task-related movement				
Aimless wandering				
Leaves, enters room				
Withdrawn (ignores others)				
Shy, timid (watches others)				
Being near, following				
Seeks reassurance, support				
Parallel play				
Work with socialization				
Collaborative work or play				
Plays verbal game				
Talks to self				
Talks to others				
Adult work (P. init.)				
Adult work (T. init.)				
Corrects others				
Pretends adult				
Copies adult				
"Me too"-copies child				
Central-telling, showing				
Central group				
"Fiddles with" materials				
Uses materials				
Uses play object as itself				
Pretends play object is something else				

	Groupings					
Task				**Social**		
Tot	1	2	3		1	2	3	Tot
				2-3 P W/A				
				2-3 P W/o A				
				4-1/2 class W/A				
				4-1/2 class W/o A				
				1/2 class + W/A				
				1/2 class + W/o A				
				P. as Ind. W/A				
				P. as Ind. W/o A	////	////	////	////

Figure 1 - Continued

Negative Affect

Tot	1	2	3	Verbal — Teacher	Non-Verbal	1	2	3	Tot
				Says "stop it" etc.	Waits for child				
				Uses threatening tone	Frowns				
				Rejects child	Ignores child				
				Blames	Points finger				
				Criticizes	Shakes finger				
				Makes threat	Pushes or pulls				
				Humiliates	Spanks				
				Yells	Other				
				Other					

Tot	1	2	3	Verbal — Pupil	Non-Verbal	1	2	3	Tot
				Says "no" etc.	Makes face				
				Teases	Frowns, pouts, withdraws				
				Laughs	Uncooperative, resistant				
				Tattles	Interferes				
				Commands or demands	Threatens				
				Makes disparaging remark	Takes prop. of another				
				Makes someone "feel small"	Damages prop. of others				
				Finds fault	Picks at child				
				Threatens	Pushes or pulls				
				Blames	Holds				
				Cries	Hits				
				Starts fight	Hurts someone with				
				Other	something				
					Other				

Positive Affect

Tot	1	2	3	Verbal — Teacher	Non-Verbal	1	2	3	Tot
				Says "thank you" etc.	Accepts favors for self				
				Agrees with child	Waits for child				
				Gives individual attention	Gives individual attention				
				Accepts favors	Sympathetic				
				Sympathetic	Listens carefully to child				
				Praises child	Smiles, laughs				
				Draws all into group	Pats, fondles, hugs child				
				Other	Other				

Tot	1	2	3	Verbal — Pupil	Non-Verbal	1	2	3	Tot
				Says "thank you" etc.	Pats, fondles, hugs toy or doll				
				Asks permission in friendly manner	Helpful, shares				
				Agrees with another	Leans close to another				
				Chooses another	Chooses another				
				Offers to compromise, share, cooperate	Smiles, laughs with another				
				Defends another	Does something for someone				
				Enthusiastic, happy	Sympathetic				
				Praises another	Pats, fondles, hugs another				
				Other	Agreeable, cooperative				
					Enthusiastic, happy				
					Other				

Figure 1 - Continued

Pupil Interest-Attention Rating Scale

Obser-vation	Interest general and high	Most pupils interested much of time	About half interested much of time	Occasional pupils interested	Pupils gen. apathetic, uninterested
1	5	4	3	2	1
2	5	4	3	2	1
3	5	4	3	2	1

Overall Emotional-Attitudinal Climate

Highly Positive	Positive most of the time	Neither Positive nor Negative	Negative Occasion-ally	Highly Negative
5	4	3	2	1
Children appear extremely happy and/or satis-fied	Most pupils appear happy and/or sat-isfied much of time	About half appear happy and/or sat-isfied much of time	Occasnlly pupils ap-pear happy and/or sat-is fied	Children appear ex-tremely un-happy and/or dissatis-fied
5	4	3	2	1

Figure 2 - Classroom Global Ratings, 1969

Program _____   School _____

Teacher _____   City _____

Grade _____   Date _____

Observer _____

Pupil Groupings

Fixed and regular for activities	Mostly fixed	Emerge about half the time, Fixed half the time	More often emerge spontaneously	Usually emerge spontaneously
1	2	3	4	5

Pupil Differentiation

Almost always work at activity	Most work at same activity most of the time	Most work at same activity half of the time	Work at different activities more often than not	Usually work at different activities
1	2	3	4	5

Teacher Congruence

Words clearly contradict evident feelings		No feelings expressed; "Accepts & Clarifies"		Completely open expression of feelings
1	2	3	4	5

Teacher Empathy

Unaware of conspicuous feeling	Occasionally aware of obvious feeling	Usually aware of obvious feeling	Sometimes aware of subtle feeling	Often aware of subtle feeling
1	2	3	4	5

Figure 2 - Continued

Pupil Reinforcement

From other pupils:	Never	Rarely	Occasionally	Frequently	Almost Constantly
	1	2	3	4	5

From adults:	Almost Constantly	Frequently	Occasionally	Rarely	Never
	5	4	3	2	1

From materials:	Never	Rarely	Occasionally	Frequently	Almost Constantly
	1	2	3	4	5

Pupil Self-Control

Rarely free
1

2

Show occasional self-control
3

4

Generally free & self-controlling
5

Show little self-control
1

Extent to which activities having clear cognitive focus characterize the classroom:

Rarely occur	About 1/4 of the time	About 1/2 of the time	About 3/4 of the time	Occur Almost constantly
1	2	3	4	5

Extent to which "game-like" activities with clear cognitive focus characterize the classroom:

Almost constantly	About 3/4 of the time	About 1/2 of the time	About 1/4 of the time	Rarely occur
5	4	3	2	1

Figure 2 A - Classroom Global Ratings, 1970

Pupil Groupings

Fixed & regular for activities	Mostly fixed	Emerge about half the time; fixed half the time	More often emerge spontaneously	Usually emerge spontaneously
1	2	3	4	5

Pupil Differentiation

Almost always work at same activity	Most work at same activity most of the time	Most work at same activity half of the time	Work at different activities more often than not	Usually work at different activities
1	2	3	4	5

Teacher Congruence

No feelings expressed	Words clearly contradict evident feelings		Some agreement of words and feelings		Words and feelings clearly agree
0	1	2	3	4	5

Teacher Empathy

Unaware of conspicuous feeling	Occasionally aware of obvious feeling	Visually aware of obvious feeling	Sometimes aware of subtle feeling	Often aware of subtle feeling
1	2	3	4	5

Pupil Reinforcement

From other pupils:	Never	Rarely	Occasionally	Frequently	Almost constantly
	1	2	3	4	5

From adults:	Almost constantly	Frequently	Occasionally	Rarely	Never
	5	4	3	2	1

From materials:	Never	Rarely	Occasionally	Frequently	Almost constantly
	1	2	3	4	5

Figure 2 A - Continued

Pupil Self-Control

Pupils: Rarely show
self-control   Occasionally show
self-control   Generally show
self-control .

| 1 | 2 | 3 | 4 | 5 |

Pupil Freedom

Pupils are: Rarely free   Occasionally free   Generally free

| 1 | 2 | 3 | 4 | 5 |

Extent to which activities having clear cognitive focus characterize the classroom:

Rarely occur	About 1/4 of the time	About 1/2 of the time	About 3/4 of the time	Occur almost constantly
1	2	3	4	5

Extent to which "game-like" activities with clear cognitive focus characterize the classroom:

Almost constantly	About 3/4 of the time	About 1/2 of the time	About 1/4 of the time	Rarely occur
5	4	3	2	1

Overall Emotional-Attitudinal Climate

Highly positive	Positive most of the time	Neither positive nor negative	Negative occasionally	Highly negative
5	4	3	2	1

Children appear extremely happy and/or satisfied	Most pupils appear happy and/or satisfied much of time	About half appear happy and/or satisfied much of time	Occasionally pupils appear happy and/or satisfied	Children appear extremely unhappy and/or dissatisfied
5	4	3	2	1

Overall, the number of items was increased, and the question of dividing the instrument in two was considered. This alternative was rejected, however, since much of the affect expressed by adults is used in classroom management, and some portion of pupil affect is interactive with adult control. The instrument is shown in Figure 3.

Teacher Practices Observation Record - The Teacher Practices Observation Record (TPOR) is an instrument developed to measure a teacher's practices in relation to John Dewey's Experimentalism (Brown, 1968). It consists of 62 sign items of teacher behavior (see Figure 4). There are no pedagogically "bad" items on the TPOR; every item describes a teacher behavior that is widely practiced in schools. However, half the items (the even numbers) describe behavior which reflects agreement with Experimentalism and would be espoused by John Dewey; the other half (the odd numbers) reflect disagreement. In the original procedure, the observer's task was to check those items which occurred during three ten-minute observation periods. The time periods used in this project have been limited to five minutes in order to parallel FLAC and FLACCS.

Brown has done extensive research with the instrument, relating it to measures of beliefs, and has shown relations between a teacher's beliefs and teaching practices, and between an observer's beliefs and what he sees in the classroom.

The TPOR provides information which relates to the instructional or pedagogical practices employed in the classroom. The major classifications of items for recording behavior are (a) Nature of the Situation, (b) Nature of the Problem, (c) Development of Ideas, (d) Use of Subject Matter, (e) Evaluation of Pupil's Work, (f) Differentiation of Tasks, and (g) Motivation and Control. Data are produced describing whether the teacher or pupil is the center of

Figure 3

Florida Climate and Control System (FLACCS), 1970[*]

Institute for the Development of Human Resources
College of Education
University of Florida
Gainesville, Florida

Program:_____   Teacher:_____

City:_____   Grade:_____

School:_____   Date:_____

Team:_____

[*]This is an experimental instrument developed by Robert S. and Ruth M. Soar, and Marjorie Ragosta, and should not be cited or used without permission.

Figure 3 - Continued

	TOT	1	2	3	Teacher
10					Teacher central
11					Leads singing, games, story
12					Moves freely among pupils
13					Leaves, enters room
14					Withdraws from class
15					Uses blackbd., A-V equip.
16					Ignores pupil
17					Attends pupil briefly
18					Attends pupil closely
19					Attends simult. activities
		1	2	3	Verbal:
20					Calls child by name
21					Directs, without reason
22					Directs, with reason
23					Questions for control
24					Calls on nonvolunteer
25					Interrupts pupil, cuts off
26					Uses time pressure
27					Bribes
28					Threatens, scolds w/conseq.
29					Gives orders, commands
30					Suggests
31					Uses sharp tone
32					Supervises pupil closely
33					Nags
34					Immobilizes pupil
35					Uses sarcasm
36					Scolds, punishes
		1	2	3	
37					Uses self report
38					Questions, states beh. rule
39					Questions for reflect. thgt.
40					Discusses behavior with P.
41					Uses humor
42					Praises
		1	2	3	Nonverbal:
43					Nods, smiles, for control
44					Tolerates deviant behavior
45					Unable to deal w/devt. beh.
46					Gesture, shhh, shakes head
47					Signals, raps, etc.
48					Takes/gives equipment, book
49					Glares, frowns
					Physical intervention:
50					Touches, pats
51					Holds, pushes
52					Punishes (other)
53					Isolates

	TOT	1	2	3	Pupil
10					Pupil Central
11					Pupil -- no choice
12					Pupil -- limited choice
13					Pupil -- free choice
					(*Seat work without teacher
					(*Seat work with teacher
16					Leaves, enters room
17					Task-related movement
18					Aimless wandering
		1	2	3	
19					Speaks aloud w/o permission
20					Engages in out-of-bound beh.
21					Collaborates with teacher
22					(*Disobeys directions
23					(*Passive resistance
24					(*Obeys directions
25					Asks permission
26					Follows routine w/o reminder
27					Reports rule to another
28					Gives direction
29					Gives reason
30					Sasses teacher
31					Tattles
		1	2	3	
32					(*Works, plays w/much supv.
33					(*Works, plays w/little supv.
34					Works, plays cooply., collab.
35					Works, plays competit.
36					Seeks reassurance, support
37					Shows pride
38					Shows shame, fear, humil.
39					Shows apathy
		1	2	3	Work Groups
40					Pupil as individual
41					Group with teacher
42					Structured groups
43					Free groups
		1	2	3	Socialization
44					Almost never
45					Occasionally, verbal
46					Frequently, verbal
47					Occasionally, nonverbal
48					Frequently, nonverbal
		1	2	3	Materials
49					Structure teacher behavior
50					Structure pupil behavior
51-52					Pupil interest-attention (Rank 1 low to 5 high)

Figure 3 - Continued

VERBAL AFFECT

TOT	1	2	3	Teacher Negative
10				Says "stop it", etc.
11				Uses threatening tone
12				Rejects child
13				Blames
14				Criticizes
15				Makes threat
16				Humiliates
17				Yells
18				Uses sarcasm
19				Other

	1	2	3	Pupil Negative
20				Says "no", etc.
21				Teases
22				Laughs
23				Tattles
24				Commands or demands
25				Makes disparaging remks.
26				Makes someone feel small
27				Finds fault
28				Threatens
29				Blames
30				Cries
31				Starts fight
32				Other

	1	2	3	Teacher Positive
33				Says "thank you", etc.
34				Agrees with child
35				Gives individual attent.
36				Accepts favors
37				Sympathetic, warm
38				Praises child
39				Draws all into group
40				Is enthusiastic, congenial

	1	2	3	Pupil Positive
41				Says "thank you", etc.
42				Sounds friendly
43				Agrees with another
44				Chooses another
45				Offers to share, coop.
46				Defends another
47				Is enthusiastic, happy
48				Praises another
49				Helps another

NONVERBAL AFFECT

TOT	1	2	3	Teacher Negative
10				Waits for child
11				Frowns
12				Ignores child
13				Points, shakes finger
14				Pushes or pulls, holds
15				Spanks
16				Shows disgust
17				Other

	1	2	3	Pupil Negative
18				Makes face
19				Frowns, pouts, withdraws
20				Uncooperative, resistant
21				Interferes
22				Threatens
23				Takes, damages property of others
24				Picks at child
25				Pushes or pulls, holds
26				Hits
27				Hurts with something
28				Is left out
29				Other

	1	2	3	Teacher Positive
30				Accepts favors for self
31				Waits for child
32				Gives individual Attention
33				Sympathetic, warm
34				Listens carefully to child
35				Smiles, laughs, nods
36				Touches, pats, hugs child
37				Other

	1	2	3	Pupil Positive
38				Helpful, shares
39				Leans close to another
40				Chooses another
41				Smiles, laughs with another
42				Sympathetic
43				Pats, hugs another
44				Agreeable, cooperative
45				Enthusiastic, happy
46				Horseplay

Ratings	
Pupil Groupings	
Pupil Differentiation	
Teacher Congruence	
Teacher Empathy	
Reinfor. from Pupils	
Reinfor. from Adults	
Reinfor. from Materials	
Pupil Self-Control	
Pupil Freedom	
Cognitive Focus	
Game-Like Activities	
Positive-Neg. Climate	
Pupils Happy, Satis.	

Figure 4 - Teacher Practices Observation Record*

TOT	I	II	III	TEACHER PRACTICES
				A. NATURE OF THE SITUATION
				1. T occupies center of attention.
				2. T makes p center of attention.
				3. T makes some thing as a thing center of p's attention.
				4. T makes doing something center of p's attention.
				5. T has p spend time waiting, watching, listening.
				6. T has p participate actively.
				7. T remains aloof or detached from p's activities.
				8. T joins or participates in p's activities.
				9. T discourages or prevents p from expressing self freely.
				10. T encourages p to express self freely.
				B. NATURE OF THE PROBLEM
				11. T organizes learning around Q posed by T.
				12. T organizes learning around p's own problem or Q.
				13. T prevents situation which causes p doubt or perplexity.
				14. T involves p in uncertain or incomplete situation.
				15. T steers p away from "hard" Q or problem.
				16. T leads p to Q or problem which "stumps" him.
				17. T emphasizes idealized, reassuring, or "pretty" aspects of topic.
				18. T emphasizes realistic, disconcerting, or "ugly" aspects of topic.
				19. T asks Q that p can answer only if he studied the lesson.
				20. T asks Q that is not readily answerable by study of lesson.
				C. DEVELOPMENT OF IDEAS
				21. T accepts only one answer as being correct.
				22. T permits p to suggest additional or alternative answers.
				23. T expects p to come up with answer T has in mind.
				24. T asks p to judge comparative value of answers or suggestions.
				25. T expects p to "know" rather than to guess answer to Q.
				26. T encourages p to guess or hypothesize about the unknown or untested.
				27. T accepts only answers or suggestions closely related to topic.
				28. T entertains even "wild" or far-fetched suggestion of p.
				29. T lets p "get by" with opinionated or stereotyped answer.
				30. T asks p to support answer or opinion with evidence.

Figure 4 - Continued

TOT	I	II	III	D. USE OF SUBJECT MATTER
				31. T collects and analyzes subject matter for p.
				32. T has p make his own collection and analysis of subject matter.
				33. T provides p with detailed facts and information.
				34. T has p find detailed facts and information on his own.
				35. T relies heavily on textbook as source of information.
				36. T makes a wide range of information material available.
				37. T accepts and uses inaccurate information.
				38. T helps p discover and correct factual errors and inaccuracies.
				39. T permits formation of misconceptions and over-generalizations.
				40. T questions misconceptions, faulty logic, unwarranted conclusions.
				E. EVALUATION
				41. T passes judgment on p's behavior or work.
				42. T withholds judgment on p's behavior or work.
				43. T stops p from going ahead with plan which T knows will fail.
				44. T encourages p to put his ideas to a test.
				45. T immediately reinforces p's answer as "right" or "wrong."
				46. T has p decide when Q has been answered satisfactorily
				47. T asks another p to give answer if one p fails to answer quickly.
				48. T asks p to evaluate his own work.
				49. T provides answer to p who seems confused or puzzled.
				50. T gives p time to sit and think, mull things over.
				F. DIFFERENTIATION
				51. T has all p working at same task at same time.
				52. T has different p working at different tasks.
				53. T holds all p responsible for certain material to be learned.
				54. T has p work independently on what concerns p.
				55. T evaluates work of all p by a set standard.
				56. T evaluates work of different p by different standards.
				G. MOTIVATION, CONTROL
				57. T motivates p with privileges, prizes, grades.
				58. T motivates p with intrinsic value of ideas or activity.
				59. T approaches subject matter in direct, business-like way.
				60. T approaches subject matter in indirect, informal way.
				61. T imposes external disciplinary control on p.
				62. T encourages self-discipline on part of p.

*Developed by Dr. Bob Burton Brown, Institute for Development of Human Resources, College of Education, University of Florida, Gainesville, Florida.

attention, the extent to which pupils are active or passive, and the amount of freedom which is permitted pupils. The nature of the problem is organized around the concerns of pupils or the concerns of the teacher or textbook, as well as the difficulty of study topics. Information is yielded as to whether ideas are treated in a "hypothetical" or "expository" manner, and whether they are dealt with in a creative or routine fashion. Subject matter is classified as to whether the pupils or the teacher assumes primary responsibility for locating it, whether it is taken from a textbook or a wide range of sources, whether it is accurate or inaccurate. Whether the teacher evaluates the pupils' work or the pupils engage in self-evaluation is recorded. The degree to which the classroom tasks are differentiated for individual pupils is measured, along with the extrinsic-intrinsic nature of the motivation and the type of the disciplinary control.

Florida Taxonomy of Cognitive Behavior - The original instrument was developed by the Florida group under Brown's leadership. Its history originates with The Taxonomy of Educational Objectives: The Cognitive Domain (Bloom and others, 1956), which was modified and extended by Sanders (1966) to provide an instrument to assess teacher lesson plans and teaching materials. The work of the Florida group has consisted of converting Sanders' instrument to one for live observation in the classroom, and of carrying out developmental work with it in classrooms. The levels into which cognitive activity is divided are as follows:

1. Memory. The student is expected to recognize or remember information. He is not expected to compare, relate, or alter the material on his own.

2. <u>Translation</u>. At this level, the student is expected to alter the form of the material with which he is dealing -- figurative to literal, behavioral to verbal, verbal to quantitative, pictorial to verbal, or abstract to concrete -- but not to change or evaluate the ideas represented.

3. <u>Interpretation</u>. The student is expected to identify similarities or differences, to compare on some other basis, to relate supporting evidence to a generalization, or to carry out a specified operation.

4. <u>Application</u>. The student is expected to bring together, without instruction, previously learned material which relates to a problem. Examples would include using word-attack skills to sound out a word, or deciding what mathematical operation is appropriate to solve a problem and carrying it through.

5. <u>Analysis</u>. This category is concerned with consciously applying the rules of thinking or of logic to the analysis of a problem, with inferring feelings or motives.

6. <u>Synthesis</u>. This level involves bringing ideas together, as in application, but with the added requirement that the student reorganizes or changes them in such a way as to produce something new. Original productions of various sorts would be classified here.

7. <u>Evaluation</u>. This level requires two functions: establishing a set of criteria, which are relevant to evaluate an idea or a product, and then evaluating the product or idea against these criteria.

In the development of the original instrument, data were collected from approximately 120 teachers using this system in parallel with the Reciprocal

Category System and the Teacher Practices Observation Record. Analysis has indicated meaningful relationships with the other instruments (Wood, 1969: Bane, 1969).

Although the initial research plan anticipated using the original form of the instrument, it was found difficult to apply to kindergarten-first grade classrooms, and a new version of the instrument was developed. The modification was developed in two stages. First, observers who had visited classrooms in the first-year data collection developed items to represent the levels of the Taxonomy from their memory of the classrooms. Then these items were tried out in tape listening, new ones developed as needed, and old ones modified or redefined. When the items stabilized, the form of the instrument was fixed and tape coding was begun (see Figure 5).

The usual recording procedure used with a sign system was modified as well. Ordinarily an item is tallied only once in an observation period, but it seemed possible that the high rate of pupil response which is emphasized in some programs might be seriously under-represented. As a consequence, the procedure of tallying every three seconds (or each interaction) was followed. Since conventional sign data typically discriminate effectively (and, in fact, were found to do so for FLAC and the TPOR in these data), the data of the Cognitive Taxonomy were also analyzed with each observation period scored zero or one (for any nonzero frequency, regardless of size). Since the latter data appeared to be at least as discriminating, only those are reported, and zero or one recording was employed with the second year's data.

A superficial consideration of the cognitive domain sometimes suggests that its higher levels are more appropriate for older pupils than younger. Yet

Figure 5 - Florida Taxonomy of Cognitive Behavior, 1969-70[*]

T	/	P	T	/	P	T	/	P		
	/			/			/		**1.**	**Memory**
	/			/			/		a.	Repeats from memory
/				/			/		b.	Repeats other
	/			/		/				
	/			/			/		c.	Repeats in sequence
	/			/			/			
	/			/			/		d.	Choral response
	/			/			/			
	/			/			/		e.	Spells
	/			/			/		f.	Gives, receives information
	/			/			/			
	/			/			/		g.	Seeks information
	/			/			/			

T	/	P	T	/	P	T	/	P		
									2.	**Translation**
	/			/			/		a.	Sounds letters
	/			/			/		b.	Names pictures, objects, color, letter
	/			/			/			
	/			/			/		c.	Copies letter, number, word (learned)
	/			/			/			
	/			/			/		d.	Gives, follows directions
	/			/			/			
	/			/			/		e.	Describes activity, picture, etc.
	/			/			/			
	/			/			/		f.	Reports experience (2+ thougths)
	/			/			/		g.	Describes situation, event
	/			/			/			
	/			/			/		h.	Recognizes word (sight reads)
	/			/			/		i.	Translates one language to another
	/			/			/		j.	Asks, gives permission

T	/	P	T	/	P	T	/	P		
									3.	**Interpretation**
	/			/			/		a.	Sounds out word
	/			/			/		b.	Classifies (1 attribute)
	/			/			/		c.	Counts
	/			/			/		d.	Adds, subtracts
	/			/			/		e.	Uses units, tens
	/			/			/		f.	Compares letters, numbers
	/			/			/		g.	Copies letter(s), number(s), learning
	/			/			/			
	/			/			/		h.	Gives class name (vehicle, etc.)
	/			/			/		i.	Identifies similarities, differences
	/			/			/			
	/			/			/		j.	Asks, gives reason
	/			/			/		k.	Names sensation
	/			/			/		l.	Performs learned task or process
	/			/			/		m.	Relates terms (of one-first)
	/			/			/		n.	Makes comparisons

Figure 5 - Continued

Activity _____   Teacher_____   Program_____

T / P	T / P	T / P	
			4. Application
/	/	/	a. Classification (2+ attributes)
/	/	/	b. Directs learning game
/	/	/	c. Creates arithmetic problem
/	/	/	d. Writes, types sentence
/	/	/	e. Asks, tells who, what, where, etc.
/	/	/	f. Seriates (alphabetizes)
/	/	/	g. Applies previous learning to new situations
/	/	/	h. Reads

T / P	T / P	T / P	
			5. Analysis
/	/	/	a. Verifies equation balance
/	/	/	b. Infers feeling or motive
/	/	/	c. Infers causality (tells why)
/	/	/	d. Cites evidence for conclusion

T / P	T / P	T / P	
			6. Synthesis
/	/	/	a. Elaborates on picture, story, etc.
/	/	/	b. Proposes plan or rule
/	/	/	c. Play acts
/	/	/	d. Makes up story
/	/	/	e. Makes fantasied object
/	/	/	f. Makes common object
/	/	/	g. Draws, colors common object
/	/	/	h. Draws, colors fantasied object

T / P	T / P	T / P	
			7. Evaluation
/	/	/	a. Compares with criteria, rule or plan

*This is an experimental form which should not be cited or used without permission of the developers.

attention to the cognitive activities of classrooms showed that activities,
at least through the middle levels of the instrument, occurred fairly frequently.
The difference, of course, is that simpler materials and concepts are involved.
The development of a Piagetian concept such as conservation falls at the level
of synthesis, and the discussion that accompanies a story or a reading lesson
may deal with questions such as, "What else might Jimmy have done?" (synthesis),
or "Would it have been better if Jimmy had done something different? Why?"
(evaluation).

The complexity of the concepts and the nature of the subject matter will
differ from age to age, of course, but higher level thought processes seem
clearly to be an important part of the development of the young child.

In fact, an idea that became more compelling as the instrument was
developed was that much of the learning done by pupils in the lower grades is
learning how to do processes that occur with little thought for older pupils.
For example, the item "Reads" is at the lowest cognitive level in the general
purpose instrument, but is a high level item for pupils at the kindergarten,
first grade level. Deriving the multiplication table is a demanding operation,
but as a tool in use it is low level, and becomes most useful when it reaches
the level of memory. Indeed, a realization that emerged which seems paradoxical
in some ways, is that a part of the process of education consists of making
higher level behaviors lower level. That is, an activity which is initially com-
plex, such as reading, becomes a lower level one as it becomes automatic and
routine. Thus, a goal of the educational process is to make complex operations
so well learned that they become low level operations, and tools in turn for
other higher level operations.

Reciprocal Category System - The work with the original system (Flanders Interaction Analysis) has been summarized by Flanders (1965, 1970), and Amidon and Hough (1967). There are a number of modifications of the system, but only the one used in this research will be discussed here.

The modification by Ober, Wood, and Roberts (1968) offers a number of advantages over the original. The seven teacher categories of the Flanders System have been expanded to nine (see Figure 6): teacher lecture is divided into that which is responsive to pupils, and that which is teacher initiated; and the category of teacher criticism has been divided into a category for correction without criticism, and one for criticism. Category 10 remains silence and confusion as before. The major advance, however, is reformulating each of the categories so that they can be used for pupil talk as well as for teacher talk. That is, teacher amplification of a pupil's idea is categorized as a 3; a pupil amplification is a 13. Each category is changed from a teacher category to a pupil category by adding a "1" as the first digit. The observer, then, learns nine categories as he did with the Flanders System but has 18 to work with and, as a consequence, the same variety of pupil talk is recorded as teacher talk. This permits identifying the extent to which pupils do such things as maintain order in the classroom, correct subject matter misunderstandings of other pupils, build on each other's ideas, contribute information, or express and accept feeling in the classroom. Practically speaking, this modification offers more than twice the richness of the data provided by the original Flanders System at little increase in the complexity of the observer's task. In the second year's data, Silence and Confusion were broken into two categories -- Silence (10) and Confusion (20).

Figure 6 - Summary of Categories for the Reciprocal Category System

Category Number Assigned to Party 1[1]	Description of Verbal Behavior	Category Number Assigned to Party 2[2]
1	<u>"WARMS" (INFORMALIZES) THE CLIMATE</u>: Tends to open up and/or eliminate the tension of the situation; praises or encourages the action, behavior, comments, ideas and/or contributions of another; jokes that release tension not at the expense of others; accepts and clarifies the feeling tone of another in a friendly manner (feelings may be positive or negative; predicting or recalling the feelings of another are included).	11
2	<u>ACCEPTS</u>: Accepts the action, behavior, comments, ideas and/or contributions of another; <u>positive reinforcement</u> of these.	12
3	<u>AMPLIFIES THE CONTRIBUTIONS OF ANOTHER</u>: Asks for clarification of, builds on, and/or develops the action, behavior, comments, ideas and/or contributions of another.	13
4	<u>ELICITS</u>: Asks a question or requests information about the content subject, or procedure being considered with the intent that another should answer (respond).	14
5	<u>RESPONDS</u>: Gives direct answer or response to questions or requests for information that are initiated by another; includes answers to one's own questions.	15
6	<u>INITIATES</u>: Presents facts, information and/or opinion concerning the content, subject, or procedures being considered that are self-initiated; expresses one's own ideas; lectures (includes rhetorical questions -- not intended to be answered).	16
7	<u>DIRECTS</u>: Gives directions, instructions, orders and/or assignments to which another is expecred to comply.	17
8	<u>CORRECTS</u>: Tells another that his answer or behavior is inappropriate or incorrect.	18
9	<u>"COOLS" (FORMALIZES) THE CLIMATE</u>: Makes statements intended to modify the behavior of another from an inappropriate to an appropriate pattern; may tend to create a certain amount of tension (i.e., bawling out someone, exercising authority in order to gain or maintain control of the situation, rejecting or criticizing the opinion or judgment of another).	19
10	<u>SILENCE</u>: Pauses, short periods of silence.	
	<u>CONFUSION</u>: Periods of confusion in which communication cannot be understood.	20

[1] Category numbers assigned to Teacher Talk when used in classroom situation.
[2] Category numbers assigned to Student Talk when used in classroom situation.

In using the Reciprocal Category System, an observer enters the class-room (or begins a tape), spends a few minutes getting the feel of what is going on, and then begins to write, at least every three seconds, the category number which best describes what is going in at that moment. If the activity changes within three seconds a new category is recorded. As a consequence the observer can sometimes record four or five categories in as many seconds. While this seems a very difficult job, eight to twelve hours of training make it relatively straightforward.

A strength of this procedure (initiated in Flanders' work) is the cap-turing, one step at a time, of the sequence of occurrences in the classroom, by the way the categories are tabulated into a matrix. It then becomes possible to answer such questions as, "What does the teacher typically do when a pupil stops talking?" "What kinds of teacher behavior are followed by pupil responses?" "Does a teacher respond differently to a pupil initiation than she does to a pupil response?" "What proportion of the teacher talk is made up of criticism of pupils. followed by directions?"

One of the interesting aspects of the matrix the RCS system produces is that it breaks down into four submatrices: teacher-teacher talk, teacher-pupil talk, pupil-teacher talk, and pupil-pupil talk. Along with this increased rich-ness of the data, the possibility is retained of returning the data to that of the Flanders System by pooling categories. As a consequence, relationships of these data to the store of information accumulated under the Flanders System can be studied easily.

The original instrument is probably the best validated of any, if validity is defined in terms of the prediction of change in pupils. The relevance of teacher behavior as measured by this instrument to pupil achievement growth has

been widely studied, and relationships with pupil attitudes have also been
found. A smaller number of researches show the validity of the instrument for
predicting such things as pupil change in personality, growth in creativity,
and perceptions of the nature of the classroom (Soar, 1966).

Observers and Training

Three teams of two observers each were used, consisting of graduate stu-
dents and research assistants. The first week of the quarter was spent in train-
ing on the Teacher Practices Observation Record (TPOR), beginning initially with
a presentation of specific categories by way of film clips, followed by observa-
tion in classrooms, with discussion of differences in the categories recorded.
The second week training was carried out on the Florida Affective Categories
(FLAC) (Florida Climate and Control System the second year), using classroom ob-
servation and discussion of differences from the beginning. The last day of the
second week all teams observed in Follow Through classrooms in Jacksonville, to
complete standardization in the use of the observation systems. The third week
all teams observed in Philadelphia, since the number of classrooms there was
large enough that all three teams could work simultaneously and could meet after
each day's observation to discuss questions raised and agree on common procedures.
This week represented the transition from training to full-scale work in the
field. In addition, some teachers seemed unusually uncomfortable about wearing
a wireless microphone. Finally, even with experience, the attention required by
the equipment continued to be a distraction of some consequence to the observers.
For these reasons, in numbers of classrooms the wireless microphones were not used.

For the second year, all of the recordings were obtained by observers moving
about as inconspicuously as possible, carrying the tape recorder. Since the re-
corders were small battery-powered units, this was easily done. Although the

distraction to the pupils was somewhat increased initially, it soon declined and seemed generally not to be a source of difficulty. The final procedure depended on choice of equipment to a large degree, and these details are presented in Appendix A.

In a free-play setting in a classroom with hard walls, floor and ceiling, it was difficult to obtain an intelligible recording. However, an observer actually present in the classroom will not be able to understand much of the interaction in such a setting. In general, the refined recording procedures seemed to produce tape which was as understandable to a coder as the live situation was to the observer.

Apart from the technical difficulties of making recordings in classrooms, there were other difficulties. The typical Follow Through classroom is an unusually diverse one in terms of the variety of activities going on simultaneously. This, in turn, means that the complexity of the observer's task is increased several fold over what it would be in the typical classroom a grade level or two higher. It seems likely that the higher the grade level, in general, the simpler the observer's task. Even with two observers watching for different classes of behaviors, it is inevitable that some portion of what occurred in the classrooms went unobserved. On the other hand, with the number of periods observed, the hope that the classroom would be fairly accurately represented seemed reasonable.

Another frequently occurring pattern of behavior made the observer's task difficult. In many classrooms, half a dozen or more small groups of pupils would be at work on different tasks, with adults with a number of the small groups. The typical behavior of the adult was to speak softly so as not to disturb other groups nearby. Pupils, to a greater or lesser degree, tended to follow this

41

same pattern. As a consequence, it was frequently difficult to hear interaction between teacher and pupils in a subgroup or to tape record it. The effect was a quiet group in a noisy room, or perhaps one not so much noisy as busy. At the extreme, but not uncommonly, it was possible to sit directly at the elbow of a teacher working with individual pupils and be unable to hear a word that was said between the teacher and the pupil.

The activities coded from tape suffer a further disadvantage in that they represent only the verbal activities in the classroom. This is not a particular problem with the Reciprocal Category System (RCS), since it is intended to record only verbal activities anyway. But for the Cognitive Taxonomy this is a somewhat greater disadvantage. A child may be working with cuisinaire rods, building a stack of blocks with a repeating sequence of colors, or carrying out a classifica- tion task by himself, and there would be no record of the cognitive complexity of this behavior recorded on the tape. To the extent that programs differ in the proportion of the learning activities that take place in verbal interaction, their representation on the Cognitive Taxonomy will differ.

On the other hand, the data collected from tape is relatively inexpensive in comparison to the cost of live data, so that it need add relatively little information to be justified.

On the whole, the data recorded live are probably reasonably representa- tive of the classrooms observed, although certainly less than complete. The data taken from tape are probably less representative, and this will need to be recog- nized in the interpretation of the results.

As another aspect of the overall procedure, data were collected on a num- ber of occasions from the Follow Through teachers in Jacksonville, toward an

analysis of the effect of the presence of an observer team over a period of
time. Before these data were analyzed, however, two studies were published
that explored the question more extensively than would have been possible here,
so the local analysis was dropped.

Observer Effects

The "conventional wisdom" of workers in this area seems to be that the
observer soon becomes part of the woodwork for the pupils, if he never inter-
acts with them and never takes part in any of the activities of the classroom
(Medley and Mitzel, 1963). He probably ceases to be a matter of concern for
the teacher much more slowly for most teachers; is never a concern for some
teachers, and probably never ceases being a concern for others.

Only recently have empirical data appeared on the question. Masling
and Stern (1969) observed two full days in each of 23 fourth and fifth grade
classrooms, and correlated observational measures at differing separations in
time from each other. They hypothesized that the effect of the observer should
diminish in time, so that later observations should correlate more highly with
each other than early ones would with late ones. They comment, "These correla-
tions show no discernible pattern over time," and conclude that two interpreta-
tions of the data are possible: ". . . (a) observer influence is negligible. . .
(b) the effects of the observer are more complex than had been foreseen and
affect various aspects of teacher and pupil behavior differntially. It is
difficult to tell from the present data which conclusion is more appropriate or
even if both cannot legitimately be made" (p. 353).

Samph (1968) made tape recordings without the teacher's knowledge, and
compared these to behavior recorded live by an observer. Teachers' agreement
to participate in a study of pupil behavior was obtained, four microphones were

installed in each classroom, and teachers were told that recording would not be begun until after pupils had had time to get used to the presence of the microphones. A month later teachers were told that recording would soon begin, but it had actually begun ten days after the microphones were installed. During this early period, control or baseline data were collected, using Flanders Interaction Analysis. Following this, observers collected live data. After the completion of data collection, teachers were informed of the deception and their permission to use the data was solicited. Teachers were also asked to indicate the style of teaching they thought ideal on the same dimensions studied in the research.

The finding of primary relevance to this study is the comparison of the baseline data to data collected when a previously scheduled observer was present in the classroom. Five variables from Flanders Interaction Analysis were tested for significance of change (all comparisons were in terms of deviations of each teacher from her own ideal). Significant change was found for two of the five variables: the amount of praise produced by the teacher increased when an observer was present, and the amount of criticism decreased. In each case the difference between means for the control and experimental conditions was about three quarters of a standard deviation. Again, this is the variability of differences between observed and ideal behavior for individual teachers, and probably is much smaller than the variability of behavior across teachers.

None of the other three variables showed significant change. They were the total of teacher acceptance of pupil's ideas, the I/D ratio (the ratio of the teacher's acceptance of feeling, praise, acceptance of ideas and questioning to her lecturing, giving directions, and criticizing), and the i/d ratio (similar to I/D, but omitting questions and lecture, the primarily substantive categories).

The changes were roughly a third of a standard deviation or less for these
latter differences.

It seems reasonable to assume that teaching is a difficult and complex
task, and that altering one's style is easiest for the more obvious aspects,
such as praising pupils more and criticizing them less. By this interpretation
the more complex measures of teacher behavior may have changed little, either
because most teachers do not teach by a conceptual scheme that includes them,
or because they are more difficult to monitor.

Overall, even the statistically significant changes do not appear great
in terms of the variability of behavior from teacher to teacher, so that it
seems reasonable to assume that teacher behavior does not change greatly as a
consequence of the presence of an observer. If the change a teacher makes is
in the direction of a truer implementation of her philosophy as Samph's study
suggests, and if programs in Follow Through follow different philosophies,
then the effect of an observer should be to sharpen program differences. But
when it is recognized that the present study is analyzing approximately 400
items of classroom behavior, it seems reasonable to hope that not many of them
were affected very much.

Analysis of Observational Data

As each five-minute observation period was completed, the nature of the
activity was described briefly on a record sheet. After the data collection
phase of the project was over, several schemes for classifying these activities
were tried out. It was immediately apparent that not all activities that
occurred could be separately classified without the data being so finely broken
down that analysis could not be carried out within any one of the categories.
The final classification scheme that was adopted was one which represented

45

simultaneously the degree of cognitive focus and of teacher direction; ranging from clear "teaching" situations, in which the objectives were clear to the teacher and to the pupils; to an intermediate level, in which the objectives were clear to the teacher, but the pupils might not have been aware that the activity had an educational purpose (story time, for instance); to activities in which the nature and direction of the activity was a matter of the pupil's choice (free-play, for example).

The conclusion was reached that analyzing separately by category was not functional on the basis of an initial factor analysis of the TPOR in which each activity by each teacher entered the analysis separately. For many teachers, then, three different sets of observational data entered the analysis, one for each activity level. As a check on the usefulness of this somewhat unconventional procedure, however, the data for all activities were pooled for each teacher, and the analysis repeated. An apparently clearer set of factors emerged, so that for all subsequent analyses all observation data for each classroom were pooled.

The next step in the analysis of each set of observational data was to calculate means and standard deviations for each of the items. Items with very small means and/or variances were then either eliminated or pooled with related items. Following this procedure, an area transformation was carried out item by item to make the data as nearly normally distributed as possible, and with approximately equal variabilities. The first of these considerations seemed important since many items showed essentially "J-curve" distributions; and equality of variances was important because incomplete factor scores were to be calculated later which would involve simply summing and averaging a series of items of behavior without further weighting.

The data for each observation system were next factor analyzed separately using principle components factor extraction with multiple R squared in the diagonal, followed by varimax rotation of a series of numbers of factors. Rather than relying on one or more of the various rule-of-thumb criteria for number of factors to rotate, a series of factors were rotated, and the output interpreted. The number of factors rotated which seemed to offer the clearest set of interpretations was retained; although on several occasions additional series of factors were rotated, seeking greater clarification. Earlier research (Soar, 1966) had suggested that the usual criteria for selecting number of factors to rotate are not functional for observational data of this sort, and the results for these analyses continue to support that conclusion. Examination of the eigenvalues, for example, shows that few factors were retained for which the eigenvalue was less than three, and even fewer for which it was less than two. When factors were rotated to eigenvalues near one, the factors seemed unreasonably fractionated, or uninterpretable.

After the decision had been reached about the number of factors to be rotated for each observation schedule, incomplete factor scores were calculated by simply pooling the T-scores for those measures which loaded $\pm.40$ or above on each of the factors. Although Glass and Maguire (1965) have criticized this procedure, Horn's comments (1965) seem more compelling. He points out that factor analysis, as any other least squares estimating procedure, will capitalize on idiosyncratic variance, and that small numbers of subjects and large numbers of measures aggravate this problem. As a consequence, validity shrinkage on cross-validation becomes extensive. The incomplete factor score procedure cited above minimizes the effects of idiosyncratic variances and validity shrinkage. It does so at the cost of permitting factors to be correlated, rather than orthogonal,

as complete factor scores would be. This is the major criticism made by Glass and Maguire, but since this assumption is typically made only for computational convenience, it seems the much less compelling issue of the two. Following the calculation of factor scores, differences between program means were tested using Duncan's new multiple range test (Dixon, 1970).

In addition to the orthogonal rotations described above, oblique rotations were also carried out with each set of data for the first year using the simple loadings procedure (Jennrich and Sampson, 1966). In no case for any set of data did two factors correlate as highly as .3, and correlations above .2 were quite rare, so that the varimax rotations were retained in all cases. (This conclusion applies to the complete rotated factors, not the incomplete factor scores).

These procedures (except for oblique rotations) were repeated for the data of the second year. Since factor structures differed to varying degrees from one year to the other, each year's structure was applied to both sets of data, and the structure retained which seemed clearest, discriminated programs best, or related most strongly to pupil growth.

In the first year's analysis of FLAC, the affect items had been summed within the eight subsections (Teacher, Nonverbal, Supportive; Pupil, Nonverbal, Supportive, etc.), and the pools had entered into analysis. Subsequent analysis of items indicated that they were not homogeneous within a subsection, so in the second year's analysis of FLACCS one factor analysis was carried out on the control items, another on the affect items, and a third on the items that had loaded most heavily on each of the factors from the first two analyses. This latter analysis is the one reported. Because of the extensive revision of the instrument, no attempt was made to report both years' data on the same set of factors.

The rating scales were also correlated with the classroom mean pupil gain

measures for each year's data. Some items are the same, but some have been re-
vised, others eliminated, and new ones added. The data are reported separately
for each year.

Analysis of Pupil Data

The first year, SRI divided several standardized and experimental tests
into parallel thirds to minimize testing time, and administered this battery in
the fall and spring to a sample of Follow Through classrooms. Pre- and post-
scores were available to this project for some of the classrooms in our sample
from these administrations. The tests used in first grade included the Metro-
politan Readiness Test, one of Deutsch's Early Childhood Inventories, and the
Caldwell-Soule Pre-School Inventory. In kindergarten, in addition to these, the
Lee-Clark Reading Readiness Test and the other Early Childhood Inventory were
given. Data were available for at least half of the pupils in 20 first grade
classrooms, and from 23 to 35 kindergarten classrooms, depending on the test.
Additional measures were administered, but not analyzed because of missing data.

These data were reduced to regressed gain scores and factor analyzed
separately by kindergarten and first grade. Factor scores were calculated from
these analyses, reduced to means by classroom, and these means were then cor-
related with factor scores from the observational data. In addition, selected
sets of data were studied by fitting curves to the relations between observa-
tional data and pupil growth data. The usefulness of this approach was sug-
gested by earlier work (Soar, 1968).

The second year, SRI changed both the test battery and the grouping of
pupils for testing, but abbreviated versions were used again. Kindergarten
classes were given the Lee-Clark Reading Readiness Test, two of the subtests
from the Deutsch Early Childhood Inventory, the Caldwell-Soule Pre-School

Inventory, the Wide Range Achievement Test, and a self-concept measure assembled from two existing measures. Entering first grade classes (those without kinder-garten experience in Follow Through) were given the same battery. Nonentering first grade pupils (those with experience in Follow Through kindergartens) were given the Metropolitan Readiness Test, two tests made of items supplied by spon-sors, the Wide Range Achievement Test, and the same self-concept measure. The achievement tests were combined into measures called "quantitative-computational," "verbal-linguistic," and "perceptual-motor."

In our analysis of the first year's data, factor analysis of the gain scores from each grade level had produced two factors, one identified as Simple-Concrete, the other as Abstract-Complex, which correlated differently with the classroom observational measures. So it seemed desirable to attempt to create new groupings of items and subtests paralleling the factors identified in the analyses of the first year's data.

Procedures paralleling those of the first year were followed in a general way, except that the need to work with regressed gain for items as well as sub-tests created problems in the analysis (the Sponsor Books, and the WRAT to some extent, included heterogeneous items not grouped into conventional subtest form, with possible scores ranging from one to twenty). As a consequence, some of the subscores in the second year's data come from factor analysis, and others come from item analysis using a priori clusters of items or results from factor analy-sis as a starting point. As before, factor scores were reduced to classroom means, and correlated with the classroom observation data. The details of the problems, the rationale, and the procedures which were evolved, are presented in Appendix B.

In the first year's data, 55 of the 70 classrooms observed are repre-sented in at least some of the pupil data analyses. Some classrooms were

observed for the sake of program description which were known not to be in SRI's

sample. In the second year's data, only 40 classrooms are represented in the

analysis of pupil data. Again, some classrooms not in SRI's sample were followed;

other reasons for loss of classrooms are detailed in Appendix C.

Results

Pupil Regressed Gain Measures

As described in the procedure section, measures of Pupil growth differing

in complexity or abstractness were sought from the measures administered by SRI,

using factor analysis and item analysis.

Achievement Data, 1969 - A two factor rotation for first grade and a

three factor rotation for kindergarten are shown in Table 3. It can be seen that

there is considerable similarity across the two analyses, despite the larger num-

ber of tests for kindergarten. In general, the Simple-Concrete measures require

performance such as naming letters or numbers, performance which seems not to re-

quire processing of information but only retrieval. The Complex-Abstract factor

score, on the other hand, requires pupils to compare complex figures, to copy

figures or shapes, or to possess broad information which is not particularly school

oriented.

Achievement Data, 1970 - In preparation for calculating regressed gain

scores for the four subgroups of pupils defined by black, white, high social

status and low social status, fall and spring means were calculated on each mea-

sure. Table 4 reports results for the a priori composites of Simple-Concrete

and Complex-Abstract pupil growth for the four subgroups. There does not appear

to be a consistent tendency for more growth in blacks or whites, or for high or

low social status to grow more during the year. The small differences that appear

Table 3

Factor Analyses of Kindergarten and First Grade
Regressed Gain Scores - 1969

Subtest Description	Kindergarten Factors[1]			h^2	First Grade Factors[2]		h^2
	1	2	3		1	2	
Lee-Clark Reading Readiness							
1. Letter Matching	.05	.62*	.28	.46	Test		
2. Letter Cross-out	.11	.48*	.28	.32	not		
3. Vocab. & Instructions	.03	.54*	-.04	.30	given		
4. Ident. Letters & Words	.24	.57*	.16	.41			
Metropolitan Readiness							
1. Word Meaning	.46*	.15	.03	.24	.49*	.03	.24
2. Listening	.38	.12	-.08	.17	.36	-.07	.14
3. Matching	.46*	.32	.09	.32	.39	.20	.19
4. Alphabet	.33	.22	.46*	.37	.24	.58*	.40
5. Numbers	.28	.31	.24	.24	.33	.29	.19
6. Copying	.33	.20	.16	.18	.44*	.14	.21
Deutsch Early Childhood Inventory							
1. Pre-Math	.39	.21	-.03	.20	Test		
2. Pre-Science	.42*	.01	.09	.18	not		
3. Pre-Positions	.43*	.00	.26	.25	given		
Deutsch Early Childhood Inventory							
1. Alphabet	.00	.08	.73*	.54	-.09	.71*	.51
2. Numerals	.10	.20	.57*	.38	.20	.62*	.42
3. Shape Names	.31	.11	.25	.17	.46*	.18	.25
Pre-School Inventory (Caldwell-Soule)							
1. Social Responsiveness	.48*	.06	.19	.28	.39	.38	.30
2. Associate Vocabulary	.54*	-.09	.25	.37	.54*	.35	.41
3. Concept Activation Numerical	.51*	.11	.34	.38	.33	.22	.16
4. Concept Activation Sensory	.27	.13	.32	.19	.42*	.13	.19
Factor Title	Complex-Abstract	Un-named	Simple-Concrete		Complex-Abstract	Simple-Concrete	

[1]N = 731

[2]N = 476

*Measures loading .40 or over were carried into factor scores.

Table 4

Fall and Spring Standings for Four Kindergarten Subgroups
on A Priori Composite Scores
1970

Subtest Description	N	Means		Raw Gain	Standard Deviations	
		Fall	Spring		Fall	Spring
Complex-Abstract Composite[1]						
Black, High SES	73	9.26	12.15	2.89	2.67	2.61
Low SES	105	9.45	11.60	2.15	3.26	3.49
White, High SES	36	12.44	14.64	2.20	3.30	2.90
Low SES	35	11.09	13.71	2.62	2.99	2.89
Simple-Concrete Composite[2]						
Black, High SES	73	12.16	17.89	5.73	4.28	4.55
Low SES	105	9.62	15.91	6.29	4.73	5.44
White, High SES	36	14.72	20.61	5.89	5.90	2.33
Low SES	35	10.71	16.94	6.23	5.68	5.32

[1]Preschool Inventory: Social Responsiveness, Associate Vocabulary, Concept Activation, Numeric

[2]Early Childhood Inventory: Alphabet, Numerals

show no consistent pattern to be associated with the particular subgroup. This conclusion was supported by tests of differences in regression coefficients for blacks and whites for a subgroup of the first year data (Honeycutt, 1971), which showed one significant difference for 13 comparisons. The t's were more often less than 1 (Table 5).

This was a convenience for the statistical analysis, but more important, the implications that it has for education seem considerable. In the current climate of controversy, one issue of which revolves around the question of whether black pupils are capable of learning abstract concepts (or whether lower class pupils are capable of such learning), the finding of such similar patterns of growth during school year is very reassuring. The black subgroups do start at a lower position and finish at a lower position than white subgroups do; and lower social status groups start and finish in lower positions than high socio-economic groups do, but growth during the year appears to be similar. These results appear to agree with those of Hayes and Grether (1969) who found that the major differences between social status groups in the amount of academic growth that took place during elementary school occurred during the summers, rather than during the school year. In their data, slopes representing growth during the school year were essentially parallel across socio-economic status groups, but they diverged during the intervals representing the summers. These data agree in general with that conclusion.

The results of the factor analyses of the kindergarten and entering first grade data are shown in Table 6. It can be seen that there is considerable similarity across the two factor analyses -- perhaps a surprising degree of similarity when it is recognized that the kindergartens were largely urban and northern, and covered a range of programs from contingency management to "open" classrooms,

Table 5

t-tests for the Regression Coefficients of Pretest on Raw Gain Scores

Test or Subtest	First Grade		t
	Black	White	
N	190	176	
Metropolitan Readiness Test			
1. Word Meaning	-.54	-.45	.94
2. Listening	-.69	-.74	.43
3. Matching	-.63	-.63	.12
4. Alphabet	-.73	-.83	3.07*
5. Numbers	-.45	-.53	1.06
6. Copying	-.44	-.47	.33
Early Childhood Inventory (Deutsch)			
1. Alphabet	-.81	-.84	.74
2. Numerals	-.73	-.73	.07
3. Shape Names	-.49	-.54	.66
Pre-School Inventory (Caldwell-Soule)			
1. Social Responsiveness	-.70	-.76	.67
2. Associate Vocabulary	-.53	-.68	1.49
3. Concept Activation, Numerical	-.53	-.54	.06
4. Concept Activation, Sensory	-.77	-.68	1.07

*$p < .01$

Table 6

Factor Analyses of Kindergarten and Entering First Grade Regressed Gain Scores - 1970

	Kindergarten Factors[1]				Entering First Factors[2]			
	1	2	3	h^2	1	2	3	h^2
Lee-Clark Reading Readiness								
1. Letter Cross-out	-.02	.04	.67*	.45	.21	.20	.49*	.32
2. Ident. of Letters & Words	.08	.21	.43*	.24	-.10	.08	.53*	.30
Deutsch Early Childhood Inventory								
1. Alphabet	.72*	.05	.14	.54	.67*	.04	.06	.45
2. Numerals	.52*	.04	.53*	.55	.54*	.09	.45*	.50
Pre-School Inventory (Caldwell-Soule)								
1. Social Responsiveness	.21	.18	.36	.21	.32	.30	.35	.32
2. Associate Vocabulary	.09	.15	.35	.16	.19	.33	.40*	.30
3. Concept Activation, Sensory	.23	.21	.53*	.37	.23	.25	.41*	.29
4. Concept Activation, Numerical	.17	.19	.44*	.26	.10	.32	.34	.23
Wide-Range Achievement Test								
1. Name Spelling	.36	.29	.35	.34	.05	.07	.55*	.31
2. Copying Marks	.25	.07	.26	.13	-.02	.28	.32	.19
3. Spelling Words from Dictation	.00	.71*	.14	.53	.19	.64*	.11	.45
4. Counting Dots	.24	.22	.14	.12	.03	.02	.32	.11
5. Oral Numbers	.50*	.02	.50*	.49	.08	.60*	.36	.49
6. Showing Fingers	.26	.37	.41*	.38	.15	.33	.71	.63
7. Which Is More?	.23	.43*	.13	.25	.22	.42*	.33	.34
8. Solving Oral Problems	.24	.56*	.17	.39	.12	.60*	.28	.45
9. Written Computation	.13	.70*	.18	.55	-.10	.67*	.35	.58
10. Recog. Two Letters in Name	.53*	.21	.14	.35	.47*	-.06	.24	.28
11. Naming 13 Letters	.74*	.07	.06	.55	.46*	.40	-.17	.39
12. Recognizing 10 Letters	.33	.11	.26	.19	.26	.11	-.03	.08
13. Word Reading Aloud	-.05	.78*	.16	.63	.08	.67*	-.01	.46
Factor Title	Simple-Concrete	Skill	Complex-Abstract		Simple-Concrete	Skill	Complex Abstract	

[1]N = 249

[2]N = 190

*Measures loading .40 or over were carried into factor scores.

whereas the entering first grades were primarily rural and/or southern, and involved only three programs, all stressing the personal-social development of children.

Comparison of Tables 3 and 6 shows extensive change in the battery of tests administered to pupils from the first to the second year, so it is not surprising to find a different factor structure. As with the first year's analyses of the kindergarten, three factors were rotated, although one of these was a different factor. Two of the factors seemed to parallel the first year's analysis: a Simple-Concrete factor, and a Complex-Abstract one. Instead of the factor made up of the Lee-Clark in the first year's data, a new third factor emerged which appeared to represent skills-learning such as reading, arithmetic computation, spelling, and comparing quantitative concepts. This latter factor apparently emerged as a consequence of the addition of the WRAT to the battery since it represents activities of this sort to a high degree.

The Abstract-Complex factor is rather different in nature the second year. Neither the Metropolitan nor the Copying Figures subtest from the Early Childhood Inventory was administered to these pupils the second year, and these tests had been important contributors to the first year's Abstract factor. Rather, two measures from the Lee-Clark, which had made a separate factor in the first year's kindergarten data, were added to the factor, perhaps because the instructions and the method of recording are complex, although the tasks do not appear to be. Also added to the factor was Numerals from the Early Childhood Inventory (which also loaded on the Concrete factor), Name Spelling (for the entering first grade), and Showing Fingers. All in all, these items seem closer to skill than the Abstract factor from the previous year.

Data reported in Appendix D show that at least the Simple-Concrete factor

rank orders programs differently than the Skill and Complex-Abstract factors do, for kindergarten pupils. The small number of classrooms should not be taken as representing programs, but the different orders for the same pupils do suggest that the measures differ in the kinds of growth they reflect. No differences which could not be assumed to be chance were found between entering first grade program groups. But with nine classrooms, and three programs, none of which is strongly skill-oriented, this is not surprising.

For the nonentering first grade pupils, composites were created by item analysis, using criterion item groups created by both a priori and factor analytic procedures. These composites are shown in Table 7. In preliminary analyses of the data, the items tended to break up into those which had been individually administered, and those which had been group administered. As a consequence, Simple-Concrete, Skill and Complex-Abstract composites were created separately for each mode of administration.

The relationships between these composites are shown in Table 8. The Group and Individual Abstract scores are not highly related, but mode of administration (or the nature of the task), appears to make less difference for the Skill and Simple-Concrete measures. The Complex-Abstract and Simple-Concrete measures relate at a low level, but Skill relates at a somewhat higher level with both of these. Overall, the pattern of correlations supports the concept of a dimension of degree of abstractness or complexity, with the three classes of measures spaced along it.

The Simple-Concrete and Skill measures seem much like those identified in the earlier two groups. The Individual Complex-Abstract seems relatively similar to that of the kindergarten-entering first data, but the Group Abstract seems more like the first year's Abstract factor, with considerable overlap with the subtests

Table 7

Items Combined Into Subscores for Nonentering First Grade, 1970

Complex-Abstract	Skill	Simple-Concrete
	Group	
MRT[1] Word meaning	MRT[1] Numbers	MRT[1] Alphabet
Matching	SIG[2] Word recognition	SIG[2] Count and write
Copying	Make sides equal	Order alphabet
	Make # lines	
	Add-Bal. Equation	
	Individual	
SII[3] Cone, Hidden Figures	SII[3] Reading Words	SII[3] Count from
Verbal opposites	Read nonwords	WRAT[4] Reading numbers
Similarities	Number reading	Name 13 letters
Inconsistencies	How many 10's	
WRAT[4] Oral problems	Add and subtract	
	WRAT[4] Spelling	
	Which is more	
	Written computation	
	Word reading	

[1]MRT - Metropolitan Readiness Test

[2]SIG - Sponsor Items (Group)

[3]SII - Sponsor Items (Individuals)

[4]WRAT - Wide Range Achievement Test

Table 8

Correlations Between Subscores for Nonentering First Grade[*]

			Variable								
		Var. 1	2	3	4	5	6	7	8	9
Group	Complex-Abstract	1								
	Skill	2	.28							
	Simple-Concrete	3	.17	.40						
Individual	Complex-Abstract	4	.37	.43	.29					
	Skill	5	.26	.67	.50	.52				
	Simple-Concrete	6	.19	.42	.68	.25	.44			
Total	Complex-Abstract	7	.87	.44	.29	.83	.43	.27		
	Skill	8	.29	.76	.44	.50	.98	.42	.45	
	Simple-Concrete	9	.20	.45	.87	-.29	.44	.98	.26	.44

[*]N = 209

of the Metropolitan that entered that factor.

The differences in the kinds of pupil attainment indicated by these measures seem important. Data reported in Appendix D indicate that when the same group of pupils in several programs is examined across the several measures, the program with the lowest mean pupil growth on the Group Complex-Abstract measure, has the highest pupil growth on the Skill measure.. That is, although all of the measures reflect intellectual growth, the nature of the growth differs enough to be capable of reflecting differences in program. The numbers of classrooms are too small to be representative, but the change from measure to measure for the same pupils seems important.

This finding points up another aspect of the test battery which seems important in relation to program evaluation. That is that when simply totaled, the WRAT has about as many possible points by itself as the rest of the tests in the battery combined. Since it appears to be predominantly skill-oriented, it has the potential of overdetermining any kind of battery total. Other tests contain some additional items reflecting skill attainment, whereas the items reflecting the broad, nonschool-oriented information in the Group Complex-Abstract measure are relatively few. The small number of these items makes separating out the differing kinds of growth important in evaluating Follow Through. In addition, the wide use and predictive power demonstrated in the past by the Metropolitan Readiness Test (where the Group Abstract measure occurs) seem important.

Self-Concept Data, 1969-1970 - For 1969., self-concept data were available for both fall and spring on so few classes in sufficient numbers to analyze. that no analysis was attempted.

In the 1970 data, adequate numbers of cases were available, but examination of the items suggested that they were too diverse to simply sum, so factor

analysis of fall and spring status scores was carried out (to minimize the
reliability problem). The loadings were marginal, however, suggesting that
individual items were insufficiently reliable to support analysis for change.
Further analysis of these data was not attempted

Observational Measures and Their Relation to Pupil Growth

The four instruments used for observation had over 400 items and measures
of behavior. Medley and Mitzel (1963) point out that single items typically do
not have high reliability, but reliability increases rapidly as items are pooled.
Since many of the items could be assumed to overlap with each other, factor anal-
ysis was used as a way of identifying clusters of behavior that tended to occur
together, and independently of other clusters. As indicated in the procedure
section, items with loading of $\pm.40$ or greater were combined into incomplete fac-
tor scores by summing algebraically, with equal weighting. These factor scores
were then used to test for differences between programs using the multiple range
test (Dixon, 1970). The homogeneous subsets reported in tables for each instru-
ment, and the nonsignificant ranges shown in Appendix.D, represent subgroups of
programs which do not differ significantly from each other. The presence of more
than one subset or nonsignificant range indicates the existence of a significant
difference between programs. These multiple range tests are presented in detail
in Appendix D.

Teacher Practices Observation Record (TPOR) - The factor structure from
the 1970 data was chosen as the one on which programs would be compared for the
two year data. The factors and their characteristics are described below. A
summary of the results of the multiple range tests is reported in Table 9, and
correlations of each factor with classroom mean pupil growth are presented in
Tables 10 and 11.

Table 9

Teacher Practices Observation Record
Multiple Range Tests of Program Discriminations[1]

| | 1969 | | 1970 | |
Factor	Homogeneous Subsets	F	Homogeneous Subsets	F
1 - Pupil Free Choice vs Teacher Structured Activity	4	9.83**	4	8.76**
2 - Experimental Teaching	3	4.25**	3	2.46*
3 - Less vs More Differentiation	2	5.67**	3	4.21**
4 - Nonexperimental Teaching	3	9.67**	4	8.67**
5 - Intrinsic Motivation in a Task Setting	2	2.24*	2	1.54
6 - Positively Focused Teaching	5	17.69**	4	11.74**

[1]N = 70 Classrooms

* $p < .05$
**$p < .01$

Table 10

Correlations of TPOR Factors With Pupil Growth

| Factor | Kindergarten | | | | | First Grade[4] | | Entering First[5] | | |
| | 1969 | | 1970 | | | 1969 | | 1970 | | |
	Abs.[1]	Concr.[2]	Abs.	Skill	Concr.	Abs.	Concr.	Abs.	Skill	Concr.
1 - Pupil Free Choice vs Teacher Structured Activity	.19	-.09	-.17	-.08	-.02	.22	-.07	.05	-.28	.21
2 - Experimental Teaching	.33	-.11	-.14	.04	.15	.18	-.07	.71*	.17	.21
3 - Less vs More Differentiation	-.09	-.13	-.00	-.17	-.12	-.33	-.02	-.22	-.48	.12
4 - Nonexperimental Teaching	-.35	.01	.47*	.36	-.03	-.28	-.02	.06	-.22	.25
5 - Intrinsic Motivation in a Task Setting	.35	.02	.52	.40	-.15	-.30	-.19	.76*	.45	.18
6 - Positively Focused Teaching	-.27	.13	.39	.30	-.12	-.29	.15	.27	.09	-.08

[1]N = 25 classrooms, p.05 = .41; p.01 = .53
[2]N = 33 classrooms, p.05 = .34; p.01 = .44
[3]N = 19 classrooms, p.05 = .46; p.01 = .58
[4]N = 20 classrooms, p.05 = .44; p.01 = .56
[5]N = 9 classrooms, p.05 = .66; p.01 = .80
*p < .05

Table 11

Correlations of Teacher Practices Observation Record Factors With Nonentering First Grade Pupil Growth (1970)[1]

Factor	Group			Individual			Total		
	Abs.	Skill	Concr.	Abs.	Skill	Concr.	Abs.	Skill	Concr.
1 - Pupil Free Choice vs Teacher Structured Activity	.14	-.11	.07	-.01	-.14	-.05	.05	-.13	-.01
2 - Experimental Teaching	.41	.25	.30	.57	.24	.07	.56	.25	.15
3 - Less vs More Differentiation	-.08	-.47	-.14	-.35	-.32	-.22	-.27	-.35	-.19
4 - Nonexperimental Teaching	-.57	-.04	-.33	-.26	.02	-.25	-.41	.01	-.28
5 - Intrinsic Motivation in a Task Setting	-.05	.63*	.43	.45	.66*	.43	.29	.66*	.44
6 - Positively Focused Teaching	-.38	.46	.24	.16	.57	.32	-.03	.55	.30

[1]N = 12 classrooms, p.05 = .58, p.01 = .71

*p < .05

(1) <u>Pupil Free Choice vs Teacher Structured Activity</u> - This factor
(Table 12) appears to contrast experimental teaching in the Deweyan sense (the
even-numbered items), with nonexperimental teaching (the odd-numbered items).
The positive pole (the loadings with no sign) appears to reflect a great deal
of pupil freedom of choice, a rich variety of materials, and much activity and
informality. All of the items appear to be ones which might be checked in a
free-play setting. Noticeably absent is much emphasis on the Nature of the
Problem and Development of Ideas.

The negative pole, in contrast, appears to represent a situation struc-
tured by the teacher, with restricted materials and little pupil activity. Again,
there is little representation of the Nature of the Problem or the Development of
Ideas. In a sense this factor seems like a free-play vs "teaching" situation,
but the teaching situation seems more like stage setting or classroom routine
than teaching.

The factor differentiates programs with an F of 9.83 for the first year's
data, and 8.76 for the second, both statistically significant. It does not
relate significantly to pupil subject matter growth for either year, however.

(2) <u>Experimental Teaching</u> - The essence of this factor appears to be
the involvement of the pupil with a problem which is uncertain or incomplete.
He is encouraged to explore, to sit and think, to suggest alternatives, but
is required to test his hypotheses and to be realistic (Table 13).

The factor differentiates programs significantly both years, with F's
of 4.25 and 2.46, but does not correlate with growth for kindergarten pupils
or first grade pupils the first year. For entering first pupils the second
year it correlated above .70 with Abstract (significant). Although not signifi-
cant for the nonentering first grades, it relates in the .40's with Group

Table 12

Teacher Practices Observation Record

Factor 1 - Pupil Free Choice vs Teacher Structured Activity

Item	Loading	Description
2	.71	T makes p center of attention
4	.71	T makes doing something center of p's attention
6	.70	T has p participate actively
8	.60	T joins or participates in p's activities
12	.82	T organizes learning around p's own problem or Q
32	.74	T has p make his own collection and analysis of subject matter
36	.72	T makes a wide range of information material available
46	.42	T has p decide when Q has been answered satisfactorily
54	.82	T has p work independently on what concerns p
60	.68	T approaches subject matter in indirect, informal way
1	-.53	T occupies center of attention
3	-.69	T makes some thing as a thing center of p's attention
5	-.68	T has p spend time waiting, watching, listening
11	-.74	T organizes learning around Q posed by T
31	-.40	T collects and analyzes subject matter for p
35	-.64	T relies heavily on textbook as source of information
53	-.46	T holds all p responsible for certain material to be learned
59	-.62	T approaches subject matter in direct, business-like way

Eigenvalue = 9.95

Table 13

Teacher Practices Observation Record

Factor 2 - Experimental Teaching

Item	Loading	Description
10	.57	T encourages p to express self freely
14	.76	T involves p in uncertain or incomplete situation
16	.71	T leads p to Q or problem which "stumps" him
18	.42	T emphasizes realistic, disconcerting, or "ugly" aspects of topic
20	.61	T asks Q that is not readily answerable by study of lesson
22	.74	T permits p to suggest additional or alternative answers
24	.62	T asks p to judge comparative value of answers or suggestions
26	.70	T encourages p to guess or hypothesize about the unknown or untested
34	.45	T has p find detailed facts and information on his own
42	.42	T withholds judgment on p's behavior or work
44	.47	T encourages p to put his ideas to a test
46	.54	T has p decide when Q has been answered satisfactorily
48	.65	T asks p to evaluate his own work
50	.49	T gives p time to sit and think, mull things over
60	.45	T approaches subject matter in indirect, informal way

Eigenvalue = 6.77

Abstract, and in the .50's with Individual and Total Abstract. It seems reasonable that a factor of this nature should relate to Abstract Growth, rather than to Skill or Concrete.

(3) <u>Less vs More Differentiation</u> - The factor title appears to be an adequate description for this factor, but with external/disciplinary control supporting the undifferentiated activity (Table 14). It discriminated significantly between programs both years. Although the correlations with growth were not significant, they were consistently negative. For nonentering first grade pupils, the relation exceeded -.40 for the Group Skill measure, and for entering first for Skill. A number of other correlations exceeded -.30. The direction of the relation associates less differentiation with less pupil growth.

(4) <u>Nonexperimental Teaching</u> - Although the items make clear that teaching is under the close control of the teacher, the control appears to be primarily one of keeping all the pupils together in what they are doing, and preventing pupils from straying in directions of their own (Table 15). Pupil plans which the teacher views as unlikely to be successful cannot be implemented, doubt and perplexity are prevented, hard questions or problems are avoided, and pupils are not permitted free expression. To a considerable degree, the factor seems to reflect a difference in orientation between programmed learning and an inquiry-orientation. In the first, learning is broken into small steps, and questions which cannot be answered correctly at the time are avoided. In the second, hard questions, exploration, and free expression are encouraged. This factor seems to be the obverse of Factor 2, Experimental Teaching, in many ways. Its flavor of closing off alternatives emphasizes the activity and involvement represented in Factor 2. Although experimental teaching is sometimes thought of as undirected, the "urging on" which characterized Factor 2 stands out sharply

Table 14

Teacher Practices Observation Record

Factor 3 - Less vs More Differentiation

Item	Loading	Description
17	.45	T emphasizes idealized, reassuring, or "pretty" aspects of topic
51	.76	T has all p working at same task at same time
55	.56	T evaluates work of all p by a set standard
61	.62	T imposes external disciplinary control on p
52	-.75	T has different p working at different tasks

Eigenvalue = 3.63

Table 15

Teacher Practices Observation Record

Factor 4 - Nonexperimental Teaching

Item	Loading	Description
9	.58	T discourages or prevents p from expressing self freely
13	.71	T prevents situation which causes p doubt or perplexity
15	.63	T steers p away from "hard" Q or problem
23	.44	T expects p to come up with answer T has in mind
25	.48	T expects p to "know" rather than to guess answer to Q
33	.41	T provides p with detailed facts and information
43	.81	T stops p from going ahead with plan which T knows will fail
57	.59	T motivates p with privileges, prizes, grades

Eigenvalue = 4.52

if one reads the first two words of each item. It discriminated significantly between programs in both years' data. The factor correlations are not all in the same direction. It correlates significantly positively with kindergarten Abstract for 1970, but generally negatively with other Abstract measures. This is difficult to resolve -- perhaps simply the variability associated with small samples.

(5) <u>Intrinsic Motivation in a Task Setting</u> - Although the factor is a diverse one, the central thread seems to be intrinsic motivation and the pupil activity that goes with it (Table 16). Although the activity is convergent in nature, the teacher sometimes withholds judgment and evaluates by different standards.

The factor differentiates programs significantly in the first year's data, but although two subgroups are created by the multiple range test in both cases, the differentiation does not create a significant F in the second year's data. -

The factor correlates above .40 with Skill growth in kindergarten, above .40 with Skill for entering first grade pupils, and above .70 for Abstract growth. For nonentering first grade pupils, all six correlations for Concrete and Skill growth are above .40. The pattern of correlations suggests that the active involvement and excitement of pupils in learning activities facilitates growth.

(6) <u>Positively Focused Teaching</u> - The flavor of the factor is that of the teacher who is central asking questions which build on previous work, accepting only answers which agree with her expectations, evaluating the pupil's answer immediately as right or wrong, and expecting the child to give evidence for his answer or helping him to find and correct his errors (Table 17).

Table 16

Teacher Practices Observation Record

Factor 5 - Intrinsic Motivation in a Task Setting

Item	Loading	Description
6	.41	T has p participate actively
23	.42	T expects p to come up with answer T has in mind
42	.48	T withholds judgment on p's behavior or work
56	.64	T evaluates work of different p by different standards
58	.70	T motivates p with intrinsic value of ideas or activity

Eigenvalue = 3.60

Table 17

Teacher Practices Observation Record

Factor 6 - Positively Focused Teaching

Item	Loading	Description
1	.42	T occupies center of attention
19	.59	T asks Q that p can answer only if he studied the lesson
21	.64	T accepts only one answer as being correct
25	.51	T expects p to "know" rather than to guess answer to Q
30	.58	T asks p to support answer or opinion with evidence
31	.58	T collects and analyzes subject matter for p
33	.40	T provides pupil with detailed facts and information
35	.47	T relies heavily on textbook as source of information
38	.62	T helps p discover and correct factual errors and inaccuracies
41	.57	T passes judgment on p's behavior or work
45	.72	T immediately reinforces p's answer as "right" or "wrong"
49	.51	T provides answer to p who seems confused or puzzled
59	.55	T approaches subject matter in direct, business-like way

Eigenvalue = 5.92

In contrast to Factor 4, this teacher is actively moving pupils along in a direction which she has determined; whereas in Factor 4, pupils are not so much being urged on as being prevented from straying in directions which seem to the teacher to be undesirable. Another contrast exists with Factor 2, in which the teacher gives pupils greater responsibility for their learning and divergent ideas are encouraged.

This factor discriminates programs more strongly than any other in the instrument for both years' data, with F's of 17.69 and 11.74. For nonentering first grade group, it relates positively to the Group Skill measure above .40, and to the Individual and Total Skill measures above .50. Although only suggestive, the factor also correlates negatively with Abstract growth in the .30's for nonentering first grade pupils, and in the .20's for 1969 kindergarten and first grade pupils.

In summary, the TPOR discriminates programs significantly for five of the six factors both years. It suggests that experimental teaching in the Deweyan sense fosters abstract growth, and that differentiation, intrinsic motivation and positively focused teaching support skill growth.

Florida Affective Categories (FLAC) - 1969

This instrument was modified extensively on the basis of the first year's work (and renamed). Because of the changes, data are not reported together for the two years as with other instruments, but separately for each year's instrument. The extent to which each factor from the instrument discriminated between programs, and correlated with pupil growth, is indicated in Table 18. Details of program comparisons are shown in Appendix D.

Table 18

Correlation of Florida Affective Categories with Pupil Growth Scores
and Multiple Range Test Program Discriminations - 1969

| Factor | Multiple Range Test[1] | | Kindergarten | | First Grade[4] | |
	Homogeneous Subsets	F	Abs.[2]	Concr.[3]	Abs.	Concr.
1 - Free Choice vs Structured Learning in Groups	3	7.60**	.10	-.08	.37	-.06
2 - Intermittent Work vs Group Singing or Games	2	1.21	-.03	.19	-.02	-.05
3 - Warm Emotional Climate	2	1.54	.14	-.12	.23	.25
4 - Teacher Negative Affect vs Pupil Work and Socialization with Adults	2	2.13	-.26	.02	.23	-.30
5 - Pupil Free to Withdraw, Seek Reassurance, Pretend, or Express Negative Affect	2	2.72*	.11	-.39*	.22	-.07
6 - Teacher Neutral Control vs Teacher Support in Task Setting	3	2.37*	-.28	-.37	-.55*	-.12

^{1}N = 70   * $p < .05$

^{2}N = 23   ** $p < .01$

^{3}N = 33

^{4}N = 20

(1) <u>Free Choice vs Structured Learning in Groups</u> - The major pole of this factor seems to represent diverse activities in the classroom in which the pupil is free to choose what he wants to do and how he wishes to do it. (See Table 19). Movement, activity, interaction with other pupils, and expression of positive affect appear to be the central theme of the factor.. While two measures seem a minimal identification of the other pole of the factor, they support each other in indicating a task group which had been specified by the teacher. This factor seems very similar to the first factor from the TPOR. The similarities of the titles reflect basic similarities in the factors, and the same programs stand at the extremes for this FLAC factor and the TPOR factor for both years' data. The finding that this factor has one of the highest F ratios for program discrimination, and separates programs into three subsets, along with a similar finding for the first factor of the TPOR, suggests that this dimension is one of the more important ones along which programs differ -- that is, that it reflects differences in program philosophies. The failure of such a powerful program discriminator to relate to growth of pupils seems an important finding.

(2) <u>Intermittent Work vs Group Singing or Games</u> - This appears to represent a pattern of occasional work interspersed with socialization;.one of individual and small group activity in contrast to total class involvement in games and songs (Table 20). It discriminates two subgroups, but does not relate to pupil growth.

(3) <u>Warm Emotional Climate</u> - All of the positive affect sums appear on this factor, plus some pupil negative affect with a lower loading (Table 21). It identifies two subgroups of programs, but the F is not significant and it does not relate to pupil growth.

Table 19

Florida Affective Categories System

Factor 1 - Free Choice vs Structured Learning in Groups

Item	Loading	Description
3	.70	Teacher moves freely
14	.47	Social group without adult
18	.56	Cognitive plan, pupil problem
21	.73	Pupil's task related movement
22	.46	Pupil's aimless wandering
30	-.68	Pupil's collaborative work or play
33	.41	Pupil talks to others
35	.42	Adult work, teacher initiated
37	.58	Pupil pretends adult
43	.46	Pupil uses materials
44	.78	Pupil uses play object as itself
45	.58	Pupil pretends play object is something else
53	.58	Pupil positive nonverbal
11	-.69	Task group with adult
12	-.47	Task group without adult

Eigenvalue = 6.74

Table 20

Florida Affective Categories System

Factor 2 - Intermittent Work vs Group Singing or Games

Item	Loading	Description
20	.55	Pupil intermittent work
29	.46	Pupil work with socialization
33	.52	Pupil talks to others
42	.57	Pupil "fiddles with" materials
43	.52	Pupil uses materials
1	-.56	Teacher leads singing, games
10	-.52	Teacher directs with song, music

Eigenvalue = 3.04

Table 21

Florida Affective Categories System

Factor 3 - Warm Emotional Climate

Item	Loading	Description
15	.57	Pupil seeks information
27	.54	Pupil seeks reassurance, support
36	.61	Pupil corrects others
40	.47	Pupil central - telling, showing
47	.48	Pupil negative verbal affect
48	.74	Teacher positive verbal
49	.79	Pupil positive verbal
52	.46	Teacher positive nonverbal
53	.56	Pupil positive nonverbal

Eigenvalue = 4.48

(4) <u>Teacher Negative Affect vs Pupil Work and Socialization with Adults</u> -
The factor seems to represent teacher negative affect in contrast to work and
socialization (Table 22). It creates two subgroups of programs, but the F is
not quite significant and it does not relate significantly to pupil growth. It
seems surprising that neither affect factor relates to the growth of pupils.

(5) <u>Pupil Free to Withdraw, Seek Reassurance, Pretend, or Express
Negative Affect</u> - This is apparently another factor which describes pupil free-
dom, but a different sort of freedom than Factor 1 (Table 23). Whereas Factor 1
appeared to describe pupil freedom in choosing and carrying out a task, this
factor suggests pupil freedom to do a variety of other things (withdraw, follow,
seek reassurance, or even express negative affect). Another thread is that of
the child fantasizing himself in an adult role and of taking responsibility for
maintaining the classroom in an adult way. While the teacher is not directing
or organizing, she is attending to several activities. The factor discriminates
significantly between programs, making two subsets. One of the correlations with
pupil growth is significant -- a negative relation with Concrete growth in kinder-
garten. In contrast, however, there is no suggestion that high standing on the
factor inhibits Abstract growth.

(6) <u>Teacher Neutral Control vs Teacher Support in Task Settings</u> - In some
ways this is a perplexing factor (Table 24). Two items warrant explanation.
Item 5, "Teacher supervises pupil closely," represents very close supervision in
which pupils carry out instructions a step at a time, under direction by the
teacher. Item 6, "Teacher immobilizes pupils," represents the intent of the
teacher to stop all movement or activity on the part of the pupils. The social
activities can be pictured as occurring fleetingly, even under teacher control.
The negative pole represents task activity and teacher positive affect.

Table 22

Florida Affective Categories System

Factor 4 - Teacher Negative Affect vs Pupil Work and Socialization
with Adults

Item	Loading	Description
6	.44	Teacher immobilizes pupil
25	.55	Pupil shy, timid
46	.65	Teacher negative verbal
50	.62	Teacher negative nonverbal
13	-.44	Social grouping with adult
19	-.53	Pupil absorbed in work

Eigenvalue = 2.94

Table 23

Florida Affective Categories System

Factor 5 - Pupil Free to Withdraw, Seek Reassurance, Pretend,
or Express Negative Affect

Item	Loading	Description
9	.47	Teacher attends simultaneous activities
22	.68	Pupil aimless wandering
24	.72	Pupil withdrawn
26	.40	Pupil being near, following
27	.47	Pupil seeks reassurance, support
32	.60	Pupil talks to self
34	.42	Adult work, pupil initiated
37	.56	Pupil pretends adult
47	.42	Pupil negative verbal
51	.56	Pupil negative nonverbal

Eigenvalue = 4.25

Table 24

Florida Affective Categories System

Factor 6 - Teacher Neutral Control vs Teacher Support in Task Settings

Item	Loading	Description
5	.58	Teacher supervises pupil
6	.47	Teacher immobilizes pupils
14	.52	Social group without adults
17	.61	Cognitive plan, teacher problem
28	.61	Pupil parallel play
16	-.56	Pupil gives information
20	-.47	Pupil intermittent work
52	-.44	Teacher positive nonverbal

Eigenvalue = 3.83

こ

The factor discriminates significantly between programs. The difference

in its nature from the control identified in Factor 1 is illustrated by the

finding that one of the programs which stresses freedom of choice and selection

of activity by pupils and which stood high on Factor 1, also stands high on

this factor which reflects a kind of close control. Conversely, one of the pro-

grams whose rationale encourages teacher control of subject matter placed at the

bottom of this factor. Even in those classrooms in which the items representing

close control occur most frequently they occur only occasionally. The infrequent

behavior occurring in classrooms in which pupil freedom is a part of program

rationale suggests that teachers may occasionally feel the need to impose order

on a classroom which has become more active and diverse than the teacher is com-

fortable with. These thoughts lead to a post hoc interpretation: although specu-

lative, perhaps the management of a classroom by a teacher can be thought of as

having two aspects -- a "structure," a set of limits or a sequence of activities

within which she organizes the work of the classroom; and "control," the indi-

vidual interactions between the teacher and the pupils intended to modify pupil

behavior. Perhaps what this interpretation suggests is that for many teachers

some minimum of structure is necessary and in its absence the teacher feels

occasional need for firm control.

The range of difficulties of implementing the various programs seems rele-

vant. It seems probable that a highly structured program which specifies a se-

quence of activities for teachers and pupils would be more easily implemented than

programs which rely more on spontaneous behavior by the teacher. It does not seem

surprising in a setting in which pupil autonomy is valued and structure minimized,

that a teacher should feel the occasional need to step in and stop activity.

Helping the teacher develop alternatives without structuring a sequence of

activities appears to be a more difficult teacher development problem.

The factor correlates negatively with several pupil growth measures, significantly for kindergarten Concrete and first grade Abstract growth. It seems possible that the strong control represented by the factor may inhibit growth, or perhaps the inconsistency of occasional strong teacher control in a relatively free context may be sufficiently unpredictable to the pupil as to minimize growth in a range of objectives.

FLAC appears to reflect a variety of ways in which teachers manage classrooms. A major component of the variation in this management is accounted for by Factor 1. An additional aspect of the control of the classroom is probably represented by Factor 1, one pole of which is negative affect expression by the teacher. It seems reasonable to assume that this affect expression will be directed primarily at behaviors which the teacher wishes to minimize. Factor 6 appears to represent occasional strong control in a generally free situation. What seems to be indicated is that the ways in which classrooms are structured and controlled by teachers are complex, and that the conceptual schemes we use may be too simple to represent them adequately.

Florida Climate and Control System (FLACCS) - 1970

As noted in the procedure section, the first year FLAC was revised extensively and organized around the concepts of control of the classroom and emotional climate. Since the instruments differ in major ways, the data are reported separately for each year. The results of the multiple range tests are reported in detail in Appendix D, and summarized briefly in Table 25. Correlation of the factors with pupil growth are reported in Tables 25 and 26

(1) Teacher Negative Control vs Teacher Warmth - This factor appears to represent the teacher who controls the classroom through the use of negative

Table 25

Correlations of Florida Climate and Control System (1970) with Pupil Growth Scores
and Multiple Range Test Program Discriminations

| | Multiple Range Test | | Kindergarten[1] | | | Entering First[2] | | |
	Homo. Subset	F	Abs.	Skill	Conc.	Abs.	Skill	Conc.
1 - Teacher Negative Control	2	3.10**	-.10	-.06	-.05	.23	-.05	-.49
2 - Pupil Freedom vs Teacher Structured Activity	4	14.50**	-.33	-.12	.11	-.42	-.11	-.02
3 - Warm Climate	3	3.43**	.42	.74**	.30	.58	.09	-.19
4 - Pupil Negative Behavior	2	2.27*	-.31	-.03	-.01	.51	.16	.01
5 - Structured Teacher Behavior with Praise	3	12.85**	.54*	.60***	-.03	-.32	-.49	.27
6 - Pupil Conforming Behavior	1	.42	-.09	-.07	.30	.10	-.09	-.40
7 - Effective Personal Communication	3	5.94**	.51*	.63**	.50*	.52	-.05	.25

[1]N = 19 classrooms; p.05 = .46; p.01 = .58

[2]N = 9 classrooms; p.05 = .66; p.01 = .80

* p < .05

** p < .01

Table 26

Correlations of Florida Climate and Control System (1970) with
Nonentering First Grade Growth Scores[1]

	Group			Individual			Total		
	Abs.	Skill	Conc.	Abs.	Skill	Conc.	Abs.	Skill	Conc.
1 - Teacher Negative Control	.05	-.05	.00	.07	.05	.03	.06	.03	.02
2 - Pupil Freedom vs Teacher Structured Activity	.11	-.20	-.14	-.17	-.21	-.18	-.07	-.21	-.17
3 - Warm Climate	-.11	.49	.11	.16	.40	.17	.07	.45	.15
4 - Pupil Negative Behavior	-.17	.04.	.10	.03	.27	.05	-.04	.22	.07
5 - Structured Teacher Behavior with Praise	-.37	.02	.20	-.31	.01	.21	-.36	.01	.21
6 - Pupil Conforming Behavior	-.55	.19	.14	-.21	.10	.06	-.36	.12	.09
7 - Effective Personal Communication	.34	.60*	.58*	.40	.52	.58*	.41	.54	.59*

[1]N = 12 classrooms; p.05 = .58; p.01 = .71

*p < .05

affect (Table 27). Pupils work under relatively close supervision, and respond with passive resistance. The factor discriminates significantly between programs but does not relate significantly to pupil growth. This seems a surprising result in terms of the concern for emotional climate which education holds.

Among the classrooms which stood high on the factor were several from big city schools which were undergoing change of administration in a context of neighborhood stress. The feeling of observers who revisited these same schools the third year of the project was that they had changed character dramatically and were again warm and friendly.

(2) <u>Pupil Freedom vs Teacher Structured Activity</u> - The central thread of the factor (Table 28) appears to be the contrast between various kinds of pupil freedom -- to choose an activity, a coworker, a group (and they work with little supervision), in contrast to activities which have been set by the teacher, and are closely supervised. There appears to be a strong parallel between this factor and Factor 1 of the 1969 FLAC, as well as the TPOR for both years.

The factor discriminates programs highly significantly, but does not relate to pupil growth. Again, this finding appears to support the comment made in connection with FLAC Factor 1, that the factors which are strongest in discriminating programs often do not relate to pupil growth.

(3) <u>Warm Climate</u> - This factor is made up principally of both teacher and pupil expression of positive affect in various ways (Table 29). These items are supplemented by the teacher listening carefully to the child, agreeing with the child, and supervising closely, but it is largely pupil behavior. The factor separates programs into three subgroups and discriminates significantly. It correlates significantly with kindergarten Skill, and at a lower level with kindergarten and entering first Abstract, and all three Skill measures for

Table 27

Florida Climate and Control System

Factor 1 - Teacher Negative Control vs Teacher Warmth

Item	Loading	Description
7	.77	T gives orders, commands
9	.77	T uses sharp tone
15	.79	T glares, frowns
21	.60	P passive resistance
25	.45	P works with much supervision
30	.45	P shows shame, fear, humility
36	.84	T uses threatening tone
37	.72	T criticizes
50	.78	T frowns
51	.63	T points, shakes finger
26	-.48	P works with little supervision
54	-.53	T gives individual attention
55	-.65	T sympathetic, warm (N.V.)

Eigenvalue = 7.92

Table 28

Florida Climate and Control System

Factor 2 - Pupil Freedom vs Teacher Structured Activity

Item	Loading	Description
17	.80	P freedom to choose activity
26	.57	P works with little supervision
31	.81	P freedom to group
35	.76	P free time
38	.47	P commands or demands
49	.41	P enthusiastic, happy (V)
60	.63	P chooses another
62	.48	P pats, hugs another
63	.42	P enthusiastic, happy (N.V.)
1	-.67	T central
3	-.45	T attends simultaneous activities
10	-.49	T supervises pupil closely
18	-.42	P seatwork w/o T
25	-.58	P works with much supervision
33	-.64	P time structured with T
43	-.45	T draws all into group

Eigenvalue = 7.11

Table 29

Florida Climate and Control System

Factor 3 - Warm Climate

Item	Loading	Description
10	.41	T supervises p closely
29	.59	P shows pride
41	.56	T agrees with child
44	.46	T is enthusiastic, congenial
46	.67	P sounds friendly
47	.65	P agrees with another
48	.51	P offers to share, cooperate
49	.49	P enthusiastic, happy (V)
56	.71	T listens carefully to child
58	.72	P helpful, shares
61	.59	P smiles, laughs with another
63	.46	P enthusiastic, happy (N.V.)

Eigenvalue = 5.26

Table 30

Florida Climate and Control System

Factor 4 - Pupil Negative Behavior

Item	Loading	Description
4	.55	T directs, with reason
14	.77	T tolerates deviant behavior
19	.43	P speaks aloud w/o permission
20	.70	P engages in out-of-bounds behavior
38	.48	P commands or demands
39	.72	P finds fault
40	.50	T says "thank you," etc.
52	.51	P picks at child
53	.57	P hits
57	.56	T touches, pats, hugs child

Eigenvalue = 4.98

nonentering first. Only one of the correlations is significant, but the pattern appears to be relatively consistent.

(4) Pupil Negative Behavior - A number of kinds of pupil negative affect expression and negative behavior appear in the factor (Table 30). Both the deviant behavior tolerated by the teacher (item 14), and the out-of-bounds be- havior of the pupils (item 20), are defined in terms of standards which the teacher attempts to maintain in the classroom, rather than in terms of observer standards. The teacher would apparently like the classroom to run differently than it does, but "puts on a good face." The factor creates two subgroups, and discriminates between programs but does not relate to pupil growth. Some of the same classrooms described in Factor 1 were also high on this one.

(5) Structured Teacher Behavior with Praise - The title reflects the strongest loadings, with additional structure in the teacher question or state- ment of a behavior rule (Table 31). The negative pole of the factor suggests a gentle sort of teacher control and related pupil behaviors, but seems too weak to warrant naming in the factor. The factor discriminates highly significantly between programs, creating three major subgroups. It also relates significantly positively with kindergarten Abstract and Skill growth. However, although not significant, there are negative correlations in the .30's for Abstract growth in the nonentering first grade, and in the .30's and .40's for the entering first grade for Abstract and Skill. This pattern of correlations suggest the possibility that this method of control on the teacher's part is functional for the growth of young children, but as they become older it is no longer supportive of Abstract growth. The failure of these correlations to reach significance makes any interpretation doubtful, even though there is considerable consistency in the correlations. Alternatively, these may be small sample variations.

Table 31

Florida Climate and Control System

Factor 5 - Structured Teacher Behavior with Praise

Item	Loading	Description
11	.51	T questions, states behavior rule
12	.57	T praises (control)
18	.41	P seatwork w/o T
32	.68	Materials structure T behavior
42	.71	T praises (affect)
8	-.40	T suggests
28	-.45	P seeks reassurance, support
48	-.41	P offers to share, coop.

Eigenvalue = 4.11

Table 32

Florida Climate and Control System

Factor 6 - Pupil Conforming Behavior

Item	Loading	Description
16	.57	T touches, pats
22	.64	P obeys directions
23	.49	P asks permission
24	.47	P follows routine w/o reminder
28	.42	P seeks reassurance, support
45	.48	P says "thank you," etc.

Eigenvalue = 2.97

(6) <u>Pupil Conforming Behavior</u> - The items seem generally to reflect a variety of conforming behaviors on the part of pupils (Table 32). The factor does not discriminate between programs; rather the F suggests that variability within programs is somewhat larger than that between programs. The factor does not correlate significantly with pupil growth, but the negative correlation with Group Abstract for nonentering first grade pupils approaches significance.

(7) <u>Effective Personal Communication</u> - The thread that runs through this factor appears to be one of real personal involvement between teacher and pupils (Table 33). Although the negative pole has only one item on it and consequently is not named, the impersonality of the teacher structuring pupils into a setting in which blackboard or audio-visual equipment serves as a teacher surrogate supports the interpretation. The factor discriminates significantly between programs, creating three subgroups. It correlates strongly and positively with pupil growth -- the most consistently of any factor in the instrument. Although some correlations are significant and others are not, with only three exceptions the correlations are relatively strong with measures of pupil growth across three subgroups of pupils. The factor appears to be a surprisingly strong one to have so few items in it. Probably it is symptomatic of a larger complex so that these items represent only a portion of the effective communication.

All in all, FLACCS appears to be sensitive to differences in programs, and probably even to periods of stress in schools; six of seven factors discriminated programs significantly, five beyond the one percent level. Some factors describe behavior graphically (Factor 1) and others which seem incomplete appear to tap subtle but apparently powerful aspects of teacher-pupil interaction (Factor 7).

Table 33

Florida Climate and Control System

Factor 7 - Effective, Personal Communication

Item	Loading	Description
13	.68	T nods, smiles, for control
43	.41	T draws all into group
44	.45	T is enthusiastic, congenial
63	.42	P enthusiastic, happy (N.V.)
2	-.46	T uses blackboard, A/V equipment

Eigenvalue = 3.02

This instrument continues a pattern cited earlier, in which the factors which discriminate programs most strongly are often not the ones that relate to pupil growth. Factors 2 and 5 discriminate programs most strongly, and both represent aspects of the structure of the classroom -- the relative balance of teacher structuring vs pupil freedom of choice, and the teacher structuring and controlling pupil behavior by the use of praise. Yet neither related consistently to pupil growth through all three subgroups. In contrast, the two factors representing supportive affect, Factors 3 and 7, do relate to measures of pupil growth across all three pupil groups -- not always significantly, but consistently and even highly. The implication, speculative to be sure, seems to be that the more objective, external, mechanical arrangements of the classroom are less important than the personal communication of warmth and support.

Reciprocal Category System - The factor structure from the 1970 data was selected as the one on which the 1969 and 1970 data would be examined. The details of program discrimination are presented in Appendix D, with a summary in Table 34. The correlations of the factors with measures of pupil growth are shown in Tables 35 and 36. The various measures are defined in Appendix E.

(1) Drill With Teacher Acceptance and Pupil Initiation vs Teacher and Pupil Initiation. The positive pole of the factor appears to describe a setting in which interaction between teacher and pupils occurs in relatively rapid-fire fashion (Table 37). It does not appear to be a typical drill session, however, in the sense that the teacher talks more, takes time to praise pupils, to accept answers, and to do these things at some length. The only pupil talk which occurs, however, follows teacher talk, much of it after a teacher question. The fact that pupils do occasionally initiate also seems to imply that the setting is not as tightly structured or rapidly occurring as the typical drill session. The negative pole of the factor is not a strong one, but reflects initiation by both

Table 34

Reciprocal Category System
Multiple Range Tests of Program Discrimination

Factor	1969[1]		1970[2]	
	Homo. Subsets	F	Homo. Subsets	F
1 - Drill with Teacher Acceptance and Pupil Initiation vs Teacher and Pupil Initiation	3	5.12**	2	5.39**
2 - Varied Pupil Directed Interaction	2	2.58*	3	4.06**
3 - Warm, Nonevaluative Teaching vs Teacher Correction	1	1.45	2	2.15
4 - Teacher Amplification in Extended Talk	1	.73	1	1.07
5 - Pupil Talk	1	.33	1	1.18
6 - Teacher Initiation with Pupil Interruption vs Teacher Question	2	3.22**	3	5.10**
7 - Teacher Direction vs Pupil Initiation	1	1.30	4	4.97**
8 - Supportive Pupil Talk	2	1.58	2	2.05

[1] $N = 70$ *$p < .05$

[2] $N = 65$ **$p < .01$

Table 35

Correlations of Reciprocal Category System With
Pupil Growth Factors

| Factor | Kindergarten | | | | | First Grade[4] 1969 | | Entering First[5] 1970 | | |
| | 1969 | | 1970 | | | | | | | |
	Abs.[1]	Conc.[2]	Abs.	Skill	Conc.[3]	Abs.	Conc.	Abs.	Skill	Conc.
1. Drill with Teacher Acceptance and Pupil Initiation vs Teacher and Pupil Initiation	.09	.13	.25	.08	-.19	-.17	.12	.09	.15	.06
2. Varied Pupil Directed Interaction	.28	-.21	-.05	.19	.17	.43	.00	.01	.30	-.18
3. Warm, Nonevaluative Teaching vs Teacher Correction	-.17	-.33	-.06	-.12	.04	-.07	.07	-.59	.02	.61
4. Teacher Amplification in Extended Talk	.18	.14	.01	-.23	-.23	.60**	-.16	-.23	.00	.25
5. Pupil Talk	-.22	-.20	.35	.39	.29	-.59**	.11	.37	.71*	-.21
6. Teacher Initiation with Pupil Interruption vs Teacher Question	-.17	-.02	-.14	-.21	.11	.42	.05	-.30	-.03	.72*
7. Teacher Direction vs Pupil Initiation	.24	.29	.11	-.17	-.24	.16	-.22	-.17	-.16	.46
8. Supportive Pupil Talk	-.11	-.22	-.04	-.09	.00	.19	-.07	-.10	-.54	.30

[1]N = 23 classrooms; p.05 = .41
[2]N = 33 classrooms; p.05 = .54
[3]N = 19 classrooms; p.05 = .46

[4]N = 20 classrooms; p.05 = .44   p.01 = .56
[5]N = 9 classrooms; p.05 = .66   p.01 = .80
*p < .05
**P < .01

Table 36

Correlations of Reciprocal Category System (1970)
With Nonentering First Grade Growth Scores

Factor	Group			Individual			Total		
	Abs.	Skill	Conc.	Abs.	Skill	Conc.	Abs.	Skill	Conc.
1. Drill with Teacher Acceptance and Pupil Initiation vs Teacher and Pupil Initiation	-.12	.41	.35	.27	.45	.28	.14	.45	.31
2. Varied Pupil Directed Interaction	-.36	-.31	-.23	-.06	-.30	-.26	.10	-.31	-.26
3. Warm, Nonevaluative Teaching vs Teacher Correction	-.28	-.08	-.12	-.48	-.25	-.08	-.45	-.23	-.09
4. Teacher Amplification in Extended Talk	.35	.16	.27	.25	.01	.20	.31	.04	.23
5. Pupil Talk	.29	.67*	.40	.60*	.47	.56	.53	.52	.51
6. Teacher Initiation with Pupil Interruption vs Teacher Question	.47	-.19	.06	-.06	-.37	.12	.14	-.33	.10
7. Teacher Direction vs Pupil Initiation	-.06	.11	.10	.11	.25	.09	.05	.23	.10
8. Supportive Pupil Talk	.04	.17	-.03	.15	.22	-.07	.12	.22	-.06

N = 12; p.05 = .58; p.01 = .71   *p < .05

Table 37

Reciprocal Category System

Factor 1 - Drill with Teacher Acceptance and Pupil Initiation vs
Teacher and Pupil Initiation

Item	Loading	Description
1	.44	Teacher warms, informalizes the climate
2	.79	Teacher accepts
4	.66	Teacher elicits
11	.72	Pupil responds
19	.62	Student response to teacher
21	.50	Teacher extended indirect
26	.75	Teacher narrow question
34	.57	Teacher-teacher flexibility
40	.46	Teacher elicit-initiate, percent
46	.67	Teacher talk
48	.76	Drill
52	.68	Pupil initiation following teacher indirect
55	.67	Teacher indirect-direct, percent
58	.58	Total number of tallies for all batches (raw)
59	.74	Total teacher talk (raw tallies) divided by number of batches
13	-.45	Pupil directs
18	-.46	Pupil initiation
51	-.54	Inquiry-drill, percent
56	-.47	Teacher initiation, percent

Eigenvalue = 9.07

teacher and pupils. Apparently the major distinction is between initiation and response. The factor discriminates significantly between programs both years, but does not relate to pupil growth.

(2) Varied Pupil Directed Interaction. This factor is labeled "varied" and "pupil directed" because the only teacher talk which appears on it either is responsive to a pupil question or is interactive with a pupil (See Table 38). Most items reflect pupil talk without teacher involvement. The varieties of pupil-talk are unusual. Pupils ask questions, initiate, direct, correct, and ask broad questions. The varieties of kinds of interaction between teacher and pupil and between pupils are high. The factor presumably reflects a relatively free situation.

The factor discriminates significantly between programs both years. It does not correlate significantly with pupil growth, but it approaches significance for 1969 first grade Abstract, and also correlates low positive for 1970 nonentering Group Abstract.

(3) Warm, Nonevaluative Teaching vs Teacher Correction. While the factor represents teacher warmth on the positive pole, the occurrence of two heavily loading items which involve acceptance as a ratio to other behavior, and the fact that the negative pole of the factor has an item which loads heavily for teacher correction, suggests that the dimension is basically one of teacher acceptance vs correction (Table 39).

The factor does not discriminate programs significantly, although two groups are created in the 1970 data, nor does it relate significantly to pupil growth. Although several moderately high correlations appear, they are inconsistent in direction, and probably represent small sample variability.

(4) Teacher Amplification in Extended Talk. The heaviest loadings reflect teacher amplification (extension and use) of pupil ideas, but the picture of extended teacher talk is supported by steady-state teacher talk, extended questions,

Table 38

Reciprocal Category System

Factor 2 - Varied Pupil Directed Interaction

Item	Loading	Description
5	.58	Teacher responds
10	.85	Pupil elicits
12	.45	Pupil initiates
13	.61	Pupil directs
14	.57	Pupil corrects
28	.49	Pupil broad question
32	.62	Pupil question, teacher question
33	.69	Pupil question, teacher response
35	.84	Teacher-pupil flexibility
36	.87	Pupil-teacher flexibility
37	.79	Pupil-pupil flexibility
38	.88	Total flexibility
54	.60	Pupil direction and criticism

Eigenvalue = 8.20

Table 39

Reciprocal Category System

Factor 3 - Warm, Nonevaluative Teaching vs Teacher Correction

Item	Loading	Description
1	.43	Teacher warms, informalizes the climate
17	.79	Teacher acceptance-rejection, percent
23	.43	Teacher revised I/D
39	.84	Teacher accept-correct, percent
8	-.74	Teacher corrects
58	-.43	Total number of tallies for all batches (raw)

Eigenvalue = 3.86

and the average length of teacher initiation (Table 40). The two minimally loading items on the negative pole are not named, but there is a suggestion that the extended questions the teacher asks are not broad questions.

The factor does not differentiate between programs. It does show a correlation with pupil growth significant at the one percent level, with 1969 first grade Abstract growth, and minimal support for this correlation for 1970 nonentering first grade Group Abstract.

(5) Pupil Talk. The positive pole of the factor is made up of various combinations of pupil-pupil talk and extended pupil talk except for the two items involving inquiry (Table 41). The rest of the factor suggests that these two items appear here because of the pupil talk component that enters inquiry, and because of the lack of drill activities. The factor does not discriminate between programs. Its correlations with pupil growth present a perplexing pattern. For 1969 first grade Abstract growth, it correlated significantly negatively at the one percent level. However, in 1970 for entering first grade it correlated positively at the five percent level with Skill growth and for nonentering first grade with Group Skill and Individual Abstract. These correlations were supported by correlations with all of the individual and total measures for the nonentering group. For two groups of pupils then, the measure appears to be positively related to Skill growth, but in a third group it is significantly negatively related to Abstract growth. It is possible that these are the occasional significant correlations which happen by chance, but this conclusion is hard to accept for several reasons. For the 1969 first grade data, Factor 4, extended teacher talk, relates positively with Abstract growth; and Factor 5, pupil talk, relates negatively to Abstract growth. In effect, the correlations, based on different factors, agree on teacher talk rather than pupil talk supporting Abstract growth. However, in the 1970 data, the correlations for

Table 40

Reciprocal Category System

Factor 4 - Teacher Amplification in Extended Talk

Item	Loading	Description
3	.74	Teacher amplifies
4	.45	Teacher elicits
5	.50	Teacher responds
23	.54	Teacher revised I/D
41	.75	Teacher amplify direct, percent
42	.66	Teacher extended question
44	.51	Steady-state teacher talk
49	.44	Average length of teacher initiation
12	-.42	Pupil initiates
27	-.41	Teacher broad question

Eigenvalue = 4.70

Table 41

Reciprocal Category System

Factor 5 - Pupil Talk

Item	Loading	Description
11	.49	Pupil responds
20	.83	Pupil-pupil talk
45	.93	Steady-state student talk
47	.63	Student talk
50	.90	Inquiry
51	.69	Inquiry-drill, percent
16	-.51	Teacher talk, percent

Eigenvalue = 4.77

103

Factor 5, argue against this interpretation where pupil talk relates strongly with all measures but Group Abstract. A possibility may be that an idiosyncrasy in the 1969 Abstract growth data could similarly affect both correlations there (the N was only 20 classrooms), but the correlation of 1970 Group Abstract does seem to differ from the others.

(6) Teacher Initiation with Pupil Interruption vs Teacher Question. This factor apparently represents extended teacher talk, much of which is initiation or lecture, but with pupils feeling free to break in (Table 42). The negative pole has teacher questioning as a common thread running through the three items. The factor discriminates between programs at the one percent level both years. It relates significantly to 1970 entering first Concrete growth, and correlates in the 40's with 1969 first grade Abstract growth, and with 1970 nonentering Group Abstract growth. Perhaps what the results for Factors 4, 5, & 6, may indicate is that there is a real relationship in the data which is obscured by the error associated with small numbers of cases and difficult to identify.

(7) Teacher Direction vs Pupil Initiation. The common thread that runs through the highest loadings for this factor is teacher direction, with an indication that the direction is given at some length and accompanied by criticism (Table 43). There is also evidence that the amount of teacher talk is high in general, and that it crosses a variety of categories. The negative pole of the factor indicates pupil initiation in two items and correction in a third.

The factor discriminates significantly between programs the second year, but does not relate to pupil growth.

(8) Supportive Pupil Talk. The factor is made up entirely of supportive pupil talk, with no teacher involvement (Table 44). Although the discrimination between programs is not significant, the multiple range test creates two homogeneous subgroups each year. It does not relate to pupil growth. It appears to be a measure

Table 42

Reciprocal Category System

Factor 6 - Teacher Initiation with Pupil Interruption vs
Teacher Question

Item	Loading	Description
6	.95	Teacher initiates
18	.41	Pupil initiation
29	.72	Pupil substantive interruption
31	.73	Total pupil interruption
43	.84	Steady-state teacher initiation
44	.58	Steady-state teacher talk
56	.79	Teacher initiation, percent
26	-.41	Teacher narrow question
40	-.80	Teacher elicit-initiate, percent
48	-.44	Drill

Eigenvalue = 6.33

Table 43

Reciprocal Category System

Factor 7 - Teacher Direction vs Pupil Initiation

Item	Loading	Description
7	.75	Teacher directs
16	.62	Teacher talk, percent
22	.84	Teacher extended direct
34	.42	Teacher-teacher flexibility
44	.51	Steady-state teacher talk
46	.53	Teacher talk
53	.85	Teacher direction and criticism
59	.49	Total teacher talk (raw tallies) divided by number of batches
12	-.47	Pupil initiates
14	-.44	Pupil corrects
18	-.45	Pupil initiation, percent

Eigenvalue = 5.62

Table 44

Reciprocal Category System

Factor 8 - Supportive Pupil Talk

Item	Loading	Description
24	.85	Pupil positive participation, percent
25	.90	Pupil revised I/D_1
57	.65	Pupil warms, accepts, amplifies

Eigenvalue = 2.83

of interest in its own right, however, in the sense that a number of programs have among their objectives fostering supportive social behavior on the part of pupils. This factor appears to be a reasonable representation of such an objective in reflecting the extent to which pupils praise each other, and accept and expand on each other's ideas in the content of a task which is real to them.

The RCS is one of the instruments coded from tape. It does not appear to be as powerful an instrument, either in discriminating programs or in relating to pupil growth, as the instruments coded live in the classroom. Probably this is not so much due to the instrument itself, or coding from tape, as to the fact that less sustained interaction occurs with pupils at these age/grade levels than with older children. Although relatively strong correlations appear, they do not do so with consistency, so that the conclusions are uncertain. To some degree this is probably small sample variability, but the nature of the underlying relations remains unclear.

<u>Cognitive Taxonomy (CogTax)</u>. For the Cognitive Taxonomy the first year factor structure was chosen for further analysis. As indicated in the Procedure section, the coding method differed from the first to the second year. The first year a tally was recorded every three seconds (category system), but each observation period was also scored zero or one as though the instrument had been coded sign system. Since that system appeared to be at least as effective in terms of both program discrimination and correlations with pupil growth, the coding the second year was done in traditional sign system fashion by coding each occurrence of each item only once in an observation period, regardless of the number of times it occurred. Details of program discrimination are reported in Appendix D, and summarized in Table 45; correlations with pupil growth are shown in Tables 46 and 47.

(1) <u>Highly Focused Learning Tasks</u>. The most heavily loading items in this factor seem to involve specific academic tasks: counting, arithmetic operations, comparing and copying letters and numbers. Lower loadings appear for classification and the sum of all the items at that level (Table 48).

The factor separated programs into two groups the first year (not significant) and three groups the second year (significant). It correlates significantly with the 1969 first grade concrete pupil growth and the 1970 nonentering first grade Individual and Total Skill Growth, with other supporting correlations.

The finding that this factor which focuses on skill oriented interaction relates to pupil Concrete and Skill learning, but not to Abstract, seems very reasonable.

(2) <u>Narrow vs Broad Answer</u>. The positive pole of the factor is made up primarily of level four behavior, application; the negative pole of level two, translation (Table 49). The level four item, "asks, tells who, what, where, etc." refers to a question such as, "What is a lid for?", and the pupil's response. In contrast, the question, "What was the boy doing?", referring to a picture, would be

Table 45

Cognitive Taxonomy
Multiple Range Tests of Program Discrimination

Factor	1969[1]		1970[2o]	
	Homo. Subsets	F	Homo. Subsets	F
1 - Highly Focused Learning Tasks	2	1.60	3	3.98**
2 - Narrow vs Broad Answer.	2	1.61	2	2.21*
3 - Moderately Focused Learning Tasks	2	2.13	2	3.15**
4 - Phonics	2	2.70*	2	2.21*
5 - Information Giving and Receiving	3	2.61*	1	.56
6 - Complex Thinking	1	.62	1	.59

[1] $N = 70.$ $*p < .05$

[2] $N = 65$ $**p < .01$

Table 46

Correlation of Cognitive Taxonomy with Pupil Growth Scores

Factor	Kindergarten 1969 Abs.[1]	Conc.[2]	1970[3] Abs.	Skill	Conc.	First Grade[4] 1969 Abs.	Conc.	Entering First[5] 1970 Abs.	Skill	Conc.
1. Highly Focused Learning Tasks	.01	-.14	.27	.35	-.21	.25	.52*	.27	-.35	-.37
2. Broad vs Narrow Answer	.23	.19	.07	-.19	-.19	-.47*	-.10	-.04	.17	.41
3. Moderately Focused Learning Tasks	-.11	-.11	.21	.12	-.53*	.44*	.38	.54	-.22	-.43
4. Phonics	.24	.08	.30	.28	-.06	-.05	.15	.50	-.42	-.03
5. Information Giving and Receiving	.00	-.09	.14	.01	-.11	-.58*	.12	.49	.00	-.01
6. Complex Thinking	.03	-.06	.12	-.08	-.40	-.43	.13	.23	-.29	-.13

[1]N = 25 classrooms; p.05 = .41; p.01 = .53
[2]N = 33 classrooms; p.05 = .34; p.01 = .44
[3]N = 19 classrooms; p.05 = .46; p.01 = .58
[4]N = 20 classrooms; p.05 = .44; p.01 = .56
[5]N = 9 classrooms; p.05 = .67; p.01 = .80
*p < .05

Table 47

Correlation of Cognitive Taxonomy (1975)
with Nonentering First Grade Growth Scores[1]

Factor	Group			Individual			Total		
	Abs.	Skill	Conc.	Abs.	Skill	Conc.	Abs.	Skill	Conc.
1. Highly Focused Learning Tasks	.15	.53	.40	.48	.62*	.44	.39	.61*	.44
2. Broad vs Narrow Answer	.28	.26	.34	.47	.25	.18	.44	.26	.24
3. Moderately Focused Learning Tasks	-.20	.35	.21	.22	.53	.28	.07	.50	.26
4. Phonics	-.64	-.20	-.03	-.51	-.21	.06	-.60	-.21	.03
5. Information Giving and Receiving	-.60*	.05	-.19	-.27	.09	-.16	-.42	.08	-.17
6. Complex Thinking	-.06	-.14	-.01	.31	.32	-.26	.19	.29	-.17

[1] N = 12 classrooms; p.05 = .58; p.01 = .71
*p < .05

Table 48

Taxonomy of Cognitive Behavior

Factor 1 - Highly Focused Learning Tasks

Item	Teacher Loading	Level	Description	Pupil Loading	Item
14	.43	3	Classifies (1 attribute) and gives class name	.47	46
15	.74	3	Counts	.69	47
16	.71	3	Adds, subtracts; uses units, tens	.65	48
17	.72	3	Compares letters, numbers; copies letters, numbers (while learning)	.72	49
23	.50	3	Sum of interpretation	.54	55
			Median	.51	66

Eigenvalue = 5.67

Table 49

Taxonomy of Cognitive Behavior

Factor 2 - Broad vs Narrow Answer

Item	Teacher Loading	Level	Description	Pupil Loading	Item
		1	Choral response	.42	34
22	.49	3	Makes comparisons	.47	54
26	.67	4	Asks, tells who, what, where, etc.	.65	58
28	.66	4	Reads	.43	60
29	.70	4	Sum of application	.62	61
		2	Seeks information	-.56	36
9	-.76	2	Describes activity, picture, etc. Reports experience. Describes situation event.	-.76	41
11	-.52	2	Asks, gives permission	-.42	43
12	-.55	2	Sum of translation	-.64	44

Eigenvalue = 7.12

a level two question, and illustrative of the negative pole of the factor. Contrary to what might be expected, the level four activities seem to be more focused on subject-matter, whereas the level two activities seem more general, and would, for example include "show and tell".

The factor separated programs into two groups both years, with the F significant the second year. It related significantly negatively with 1969 first grade Abstract growth, but as strongly positively, although not significantly, with nonentering first grade 1970 Individual Abstract growth. The direction of the relationship the first year indicated that higher level activities in the classroom were detrimental to Abstract growth and simpler activities were supportive of this growth, but the opposite was indicated the second year. A possible resolution of this conflict is that the first grades for 1969 were primarily entering first grades. The pattern of correlations suggests that the complex classroom activities which were inimical to Abstract growth in 1969 were supportive of such growth among nonentering first grade pupils, and that their greater experience in school laid a background to permit this. But the shifting relationships may, of course, be due to small samples, a change in the pupil measures, or perhaps the change in the way tapes were coded.

(3) <u>Moderately Focused Learning Tasks</u>. This factor overlaps with Factor 1 in kinds of activities included (Table 50). It is task-oriented, but appears to be less specifically focused. Whereas Factor 1 was focused on tasks such as counting, arithmetic operations, and the comparison of letters and numbers, Factor 3 has its heaviest loadings for the identification of more general similarities and differences, the sum of level three activities (interpretation) and the median level of activity for the whole scale.

The factor separated programs into two subgroups each year, but the F's were not significant. The factor correlated negatively with both 1970 kindergarten Concrete (significant) and with 1970 entering first grade Concrete (above .40).

Table 50

Taxonomy of Cognitive Behavior

Factor 3 - Moderately Focused Learning Tasks

Teacher				Pupil	
Item	Loading	Level	Description	Loading	Item
7	.52	2	Names pictures, objects, color, letter	.55	39
11	.46	2	Asks, gives permission		
12	.49	2	Sum of translation	.43	44
13	.55	3	Sounds out word	.54	45
14	.60	3	Classifies (1 attribute), gives class name (vehicle, etc.)	.61	46
18	.71	3	Identifies similarities, differences	.70	50
19	.61	3	Asks, gives reason	.64	51
21	.43	3	Performs learned task or process		
23	.79	3	Sum of interpretation	.75	55
65	.79		Teacher -- Median -- Pupil	.61	66

Eigenvalue = 8.37

Table 51

Taxonomy of Cognitive Behavior

Factor 4 - Phonics vs Classification

Teacher				Pupil	
Item	Loading	Level	Description	Loading	Item
6	.65	2	Sounds letters	.75	38
13	.54	3	Sounds out word	.52	45
32	.41	7	Sum of evaluation		
24	-.46	4	Classification (2 attributes)	-.48	56

Eigenvalue = 4.20

It also correlated significantly positively with 1969 first grade Abstract, above .50 for 1970 entering first Abstract and nonentering first grade Individual Skill. In this case, then, an activity which is significantly positively correlated with Abstract growth is significantly negatively correlated with concrete growth with supporting non-significant correlations. The directions and sizes of the correlations are consistent across pupil groups. There is also a suggestion that a classroom process which supports Abstract growth for less educationally experienced children shifts to supporting Skill growth for more educationally experienced pupils.

(4) Phonics. The factor is a minimal one. Two items represent phonics work, and another the teacher evaluating or seeking evaluation, versus a minimal loading for classification on the other pole. (See Table 51).

The factor discriminates significantly between programs both years. For 1970 nonentering first grade pupils, it correlates significantly negatively with Group Abstract, and above .50 negative, for Individual Abstract. It also correlates above .50 positively but not significantly with entering first Abstract, which is troublesome. (This group has the smallest number of classrooms of any.)

(5) Information Giving and Receiving. This is a factor which appears to represent exchange of information between teacher and pupil and simple translation of language into behavior (Table 52).

It discriminates significantly between programs the first year but not the second. It correlates significantly with 1969 first grade Abstract and in the .40's with 1970 entering first Abstract; but significantly negatively with 1970 nonentering first grade Group Abstract. The direction of the two sets of correlations may mean that the lower cognitive level activities which support Abstract growth for first graders without previous school experience has become too simple and no longer fosters Abstract growth for first graders who have previous school experience. Even though these correlations are significant, other reversals raise continuing questions.

(6) <u>Complex Thinking</u>. The factor is concerned principally with the application of previous learning to a new situation including directing a learning game. Other items seem to support these activities (Table 53).

The factor does not discriminate programs either year, and the F suggests that variability within programs is at least as large as between programs. The factor does not correlate with pupil growth significantly nor in a consistent manner.

The data from this instrument and from the RCS which preceded it seem preplexing. The data from both of these instruments show parallel contradictions from pupil group to pupil group in whether low cognitive level or tightly structured classroom activities relate to a given kind of pupil growth. Similar kinds of contradictions appear from 1969 first grade data to 1970 nonentering first grade data. These are in contrast to the results from the two instruments used in live observation which followed expectation more consistently.

Table 52

Taxonomy of Cognitive Behavior

Factor 5 - Information Giving and Receiving

Teacher				Pupil	
Item	Loading	Level	Description	Loading	Item
3	.80	1	Gives, receives information	.83	35
4	.77	1	Seeks information		
5	.84	1	Sum of memory	.54	37
8	.48	2	Gives, follows directions		
10	.48	2	Recognizes-word (sight reads)	.46	42
11	.41	2	Asks, gives permission		
12	.40	2	Sum of translation		

Eigenvalue = 6.10

Table 53

Taxonomy of Cognitive Behavior

Factor 6 - Complex Thinking

Teacher				Pupil	
Item	Loading	Level	Description	Loading	Item
27	.52	4	Applies previous learning to new situations	.69	59
29	.42	4	Sum of application	.64	61
30	.55	5	Sum of analysis	.49	62
		2	Gives, follows directions	.47	40
		4	Directs learning game	.45	57

Eigenvalue = 4.51

Relations of Ratings With Pupil Growth Measures.   The relationships of .
ratings with the growth of various pupil groups are shown in Tables 54, 55 and
56.   The relations taken as a whole are so inconsistent as to be of uncertain
meaning.   Several of the ratings seem to relate strongly for nonentering first
grade pupils, but do not relate for the other groups.   Teacher congruence, for .
example, relates relatively consistently, and in two cases significantly, with
pupil growth for nonentering first, but does not relate at all for the other
four groups.   Reinforcement from adults relates fairly consistently with growth
(although not significantly) for nonentering first and entering first, but not
at all for the other three groups.   Reinforcement from materials relates strongly
and often significantly with growth for the nonentering first grade group, to a
essentially zero for entering first, but negatively for 1970 kindergarten Concrete
and not at all for the other two groups.   Clear Cognitive Focus relates positively
for at least some measures in the three 1970 groups, but not for the other two.
This measure seems to relate with at least some measures in more groups than any
other.   For other measures, relationships appear to scatter sufficiently that
conclusions are uncertain.

, Relations involving 1970 nonentering first grade stand out as different,
but it seems unlikely that the difference between those ratings and the 1970
kindergarten and entering first results can be attributed to observer differences,
since the observers were spread across all three groups.   Apparently, in this case
as with the systematic observation data, the need is for more classrooms, so that
the results can be held with some confidence.

Table 54

Correlations of Global Ratings with Pupil Growth - 1969

Rating	Kindergarten Abstract[1]	Concrete[2]	First Grade[3] Abstract	Concrete
Proportion of Interested Pupils	.34	.11	.06	.17
Intensity of Interest	.28	.18	-.43	-.09
Pupil Groupings	.02	-.20	.06	.12
Pupil Differentiation	.08	.04	.21	.22
Teacher Congruence	-.01	-.02	-.15	.02
Teacher Empathy	.17	.13	-.17	.00
Pupil Reinforcement from Pupils	.08	.09	.04	.11
Pupil Reinforcement from Adults	.06	.09	-.32	.14
Pupil Reinforcement from Materials	.08	.15	-.01	.28
Pupil Self Control	.05	.05	.28	.11
Clear Cognitive Focus	.12	.20	-.26	.25
Game-like Cognitive Focus	.14	.23	.29	.23
Children's Original Art Work	.44*	.01	.25	-.14
Room Displays and Artifacts	.22	-.18	.27	.36
Overall Emotional - Attitudinal Climate	.40	.15	-.15	.27
Children's Happiness - Satisfaction	.38	.23	-.01	.13

[1] N=23 classrooms, p.05=.41   *p .05

[2] N=33 classrooms, p.05=.34

[3] N=20 classrooms, p.05=.44

Table 55

Correlations of Global Ratings with Pupil Growth - 1970

Rating	Abs.	Kindergarten[1] Skill	Conc.	Abs.	Entering First[2] Skill	Conc.
Pupil Groupings	-.44	-.24	-.24	.00	.21	-.45
Pupil Differentiation	.12	.35	-.07	-.40	.08	.43
Teacher Congruence	-.12	.03	-.22	.36	.11	-.12
Teacher Empathy	-.13	.02	-.03	.21	.26	.58
Pupil Reinforcement from Pupils	-.16	.02	.15	-.08	.23	.11
Pupil Reinforcement from Adults	.07	.24	-.11	.51	.42	.23
Pupil Reinforcement from Materials	-.21	-.17	-.42	.37	.00	.29
Pupil Self-Control	.02	-.16	-.02	-.04	-.14	.64
Pupil Freedom to Interact	-.42	-.17	-.18	-.15	.03	-.18
Clear Cognitive Focus	.51*	.25	.05	.01	-.07	.71*
Game-like Cognitive Focus	.11	.15	-.09	-.06	-.15	.38
Overall Emotional - Attitudinal Climate	.33	.23	.10	.04	.12	.33
Children Happiness - Satisfaction	.07	.09	.02	.00	.31	.47

[1]N= 19 classrooms, p.05=.46   *p .05

[2]N= 9 classrooms, p.05=.67

Table 56

Correlations of Global Ratings with Nonentering First
Grade Pupil Growth (1970)[1]

Factor	Group			Individual			Total		
	Abs.	Skill	Conc.	Abs.	Skill	Conc.	Abs.	Skill	Conc.
Pupil Groupings	.12	.21	.30	.23	.24	.12	.21	.24	.18
Pupil Differentiation	.34	.39	.47	.29	.09	.55	.34	.15	.53
Teacher Congruence	-.03	.62*	.58*	.35	.49	.54	.24	.52	.56
Teacher Empathy	.05	.33	.32	.40	.32	.15	.30	.32	.21
Pupil Reinforcement from Pupils	.21	.08	.24	.07	.02	.17	.13	.03	.19
Pupil Reinforcement from Adults	-.45	.53	.38	.17	.56	.46	-.06	.57	.44
Pupil Reinforcement from Materials	.13	.82**	.74**	.56	.66*	.77**	.44	.70*	.77**
Pupil Self Control	-.00	.41	.04	.18	.28	.09	.13	.32	.08
Pupil Freedom to Interact	.09	.10	.28	.08	.12	.13	.09	.11	.18
Clear Cognitive Focus	-.06	.81**	.53	.52	.83**	.54	.34	.83**	.54
Game-like Cognitive Focus	.28	.36	.09	.53	.55	.08	.48	.52	.08
Over All Emotional - Attitudinal Climate	.25	.68*	.45	.45	.50	.45	.41	.54	.46
Children Happiness - Satisfaction	.30	.72**	.71**	.50	.55	.67*	.46	.60*	.70*

[1]N = 12 classrooms, p.05=.58; p.01=.71

*p<.05
**p<.01

Nonlinear Relations Between Observational Measures and Pupil Growth.
Earlier work had suggested that relations between classroom behavior measures
and pupil growth measures were sometimes nonlinear, indicating that intermediate
levels of some kinds of behavior were most functional for the growth of pupils
(Soar, 1968). Further, this work had indicated that the level of a given class-
room behavior which was most supportive of pupil growth varied with the nature
(the complexity or abstractness) of the learning task. In line with this, curves
were fitted to the relations between the Abstract and Concrete pupil growth
measures and all of the classroom behavior measures for the 1969 first grade data.
(First grade was selected because relationships seemed stronger there.) One of
the findings from that analysis is presented here as illustrative of the kinds
of relationships found, but numbers of others followed the same pattern. Table
57 presents the loadings for Factor 1 from the 1969 factor structure of the TPOR,
and Figure 7 presents the curves. The deviations from linearity are not signi-
ficant (N=20), but the curve for Complex-Abstract growth suggests that the true
relation is not linear. Numbers of other curves support that hypothesis.

For the 1970 data, Figure 8 presents the relations between FLACCS Factor
6, Pupil Conforming Behavior (Table 24), and Abstract, Skill, and Concrete growth
measures for the nonentering first grade. The suggestion of nonlinearity, al-
though nonsignificant (N=12), again is present. Additional analyses of 1970
data are not reported, since nonlinear analyses are more subject to error due to
small samples than linear analyses. These findings are not presented as more
than suggestive, but further analyses will be carried out with additional data
for 1971.

Table 57

Teacher Practices Observation Record (1969)

Factor 1 - Teacher Directed Activity vs. Pupil Selected Activity

Item	Loading	Description
1	.80	T occupies center of attention
3	.72	T makes some thing as a thing-center of p's attention
5	.45	T has p spend time waiting, watching. listening
11	.70	T organizes learning around Q posed by T
13	.56	T prevents situation which causes p doubt or perplexity
19	.52	T asks Q that p can answer only if he studied the lesson.
21	.84	T accepts only one answer as being correct
23	.73	T expects p to come up with answer T has in mind
25	.55	T expects p to "know" rather than to guess answer to Q
27	.42	T accepts only answers or suggestions closely related to topic.
31	.80	T collects and analyzes subject matter for p
33	.81	T provides p with detailed facts and information
35	.81	T relies heavily on textbook as source of information
45	.73	T immediately reinforces p's answer as "right" or "wrong"
51	.52	T has all p working at same task at same time
53	.49	T holds all p responsible for certain material to be learned
55	.49	T evaluates work of all p by a set standard
57	.63	T motivates p with privileges, prizes, grades
59	.66	T approaches subject matter in direct, business-like way
4	-.50	T makes doing something center of p's attention.
6	-.40	T has p participate actively.
12	-.73	T organizes learning around p's own problem or Q
32	-.72	T has p make his own collection and analysis of subject matter
36	-.53	T makes a wide range of information material available
44	-.44	T encourages p to put his ideas to a test
54	-.71	T has p work independently on what concerns p
60	-.68	T approaches subject matter in indirect, informal way

Eigenvalue = 12.81

Simple-
Concrete
Growth

Complex-
Abstract
Growth

Pupil
Growth

50   49   48   47   46   45

42   44   46   48   50   52   54   56   58   60   62   64   66   68

Pupil Selected Activity   Factor 1   Teacher Directed Activity
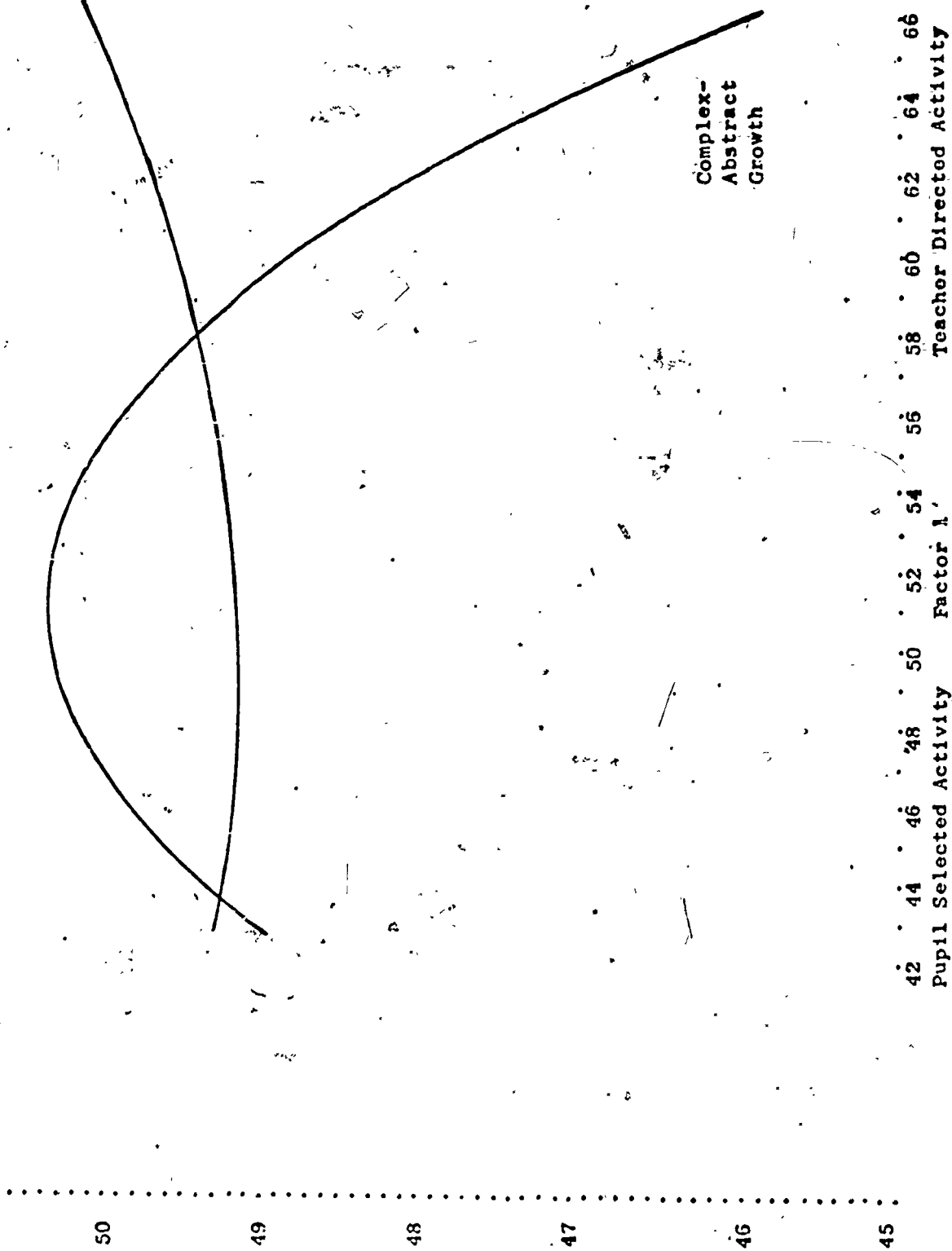
Figure 7 - Relation Between Teacher Practices Observation Record Factor 1 and Pupil Growth.
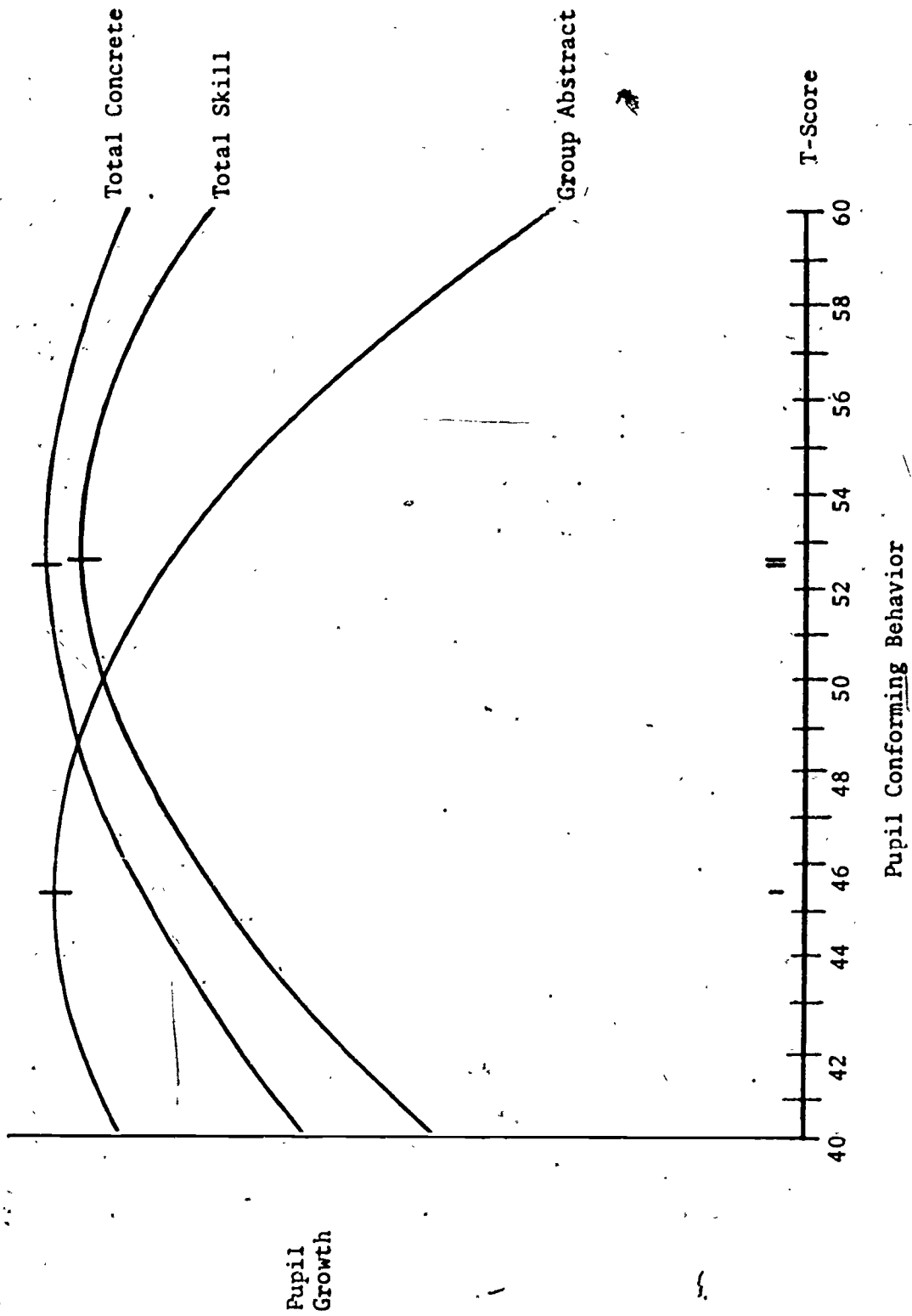
Figure 8 - Florida Climate and Control System. Factor 6 Related to Nonentering First Grade Growth.

If, however, the relations of some of the behavior measures to pupil growth are nonlinear, the correlations reported here will be underestimated. Further, where the relationships are shaped like inverted "U's", as some seem to be, the magnitude, and even the sign of a product-moment correlation may be changed with a change in the average level of the classroom behavior for different subgroups being analysed. For some measures, it is not unusual to find a change of a standard deviation or more in mean level of behavior between entering first grade, and nonentering first grade, for instance. It seems possible, then, that some portion of the variability of correlations observed may be due to different segments of a curve being represented, as well as by small sample variability.

Discussion

Program Discrimination

One of the goals of this project was to describe differences among the programs in terms of observed classroom behavior. The analyses of variance for observational data show that highly significant differences in classroom behavior are associated with differences in sponsorship. A major dimension of programs which runs through three different instruments over two years of data collection with considerable consistency separates programs so widely that there is no overlap between programs at the extremes. Numbers of dimensions which are less clearly related to sponsor objectives as we understand them also discriminate, although less sharply. And of course, some factors show no differences, but they are typically dimensions which seem to be unrelated to sponsor objectives.

There appear to be differences in the degree to which sponsors have been successful in implementing their objectives, and it seems probable that stresses in communities, which are reflected in the schools, have negated the sponsors

efforts in some cases. It also seems likely that programs differ in the dif-
ficulty of implementation. Nevertheless, the success of sponsors in producing
classrooms which reflect their objectives as they are measured here seems
striking.

Pupil Growth Data

Although not an objective anticipated as part of this project, reduction
of pupil achievement data to a small set of more homogeneous measures became a
necessary first step in preparation for studying the relations between classroom
behavior measures and pupil growth.

Probably the most important finding is that the total battery can be broken
down into largely independent subscores which represent rather different kinds
of learning. What has been called Simple-Concrete learning seems to require little
but memory; Skill represents the acquisition of the traditional academic skills,
reading, spelling and arithmetic; Complex-Abstract learning seems to require more
complex information processing -- solving complex problems, comparing complex
figures, the possession of abstract information. Not only do these different
classes of scores not relate strongly with each other, but they seem to respond
differently to the dimensions of classroom behavior and to different programs.
For example, when the pupils in several programs are examined across the three
kinds of measures, one program goes from the lowest standing to the highest
standing.

The finding of relatively independent classes of measures seems especially
important since the total battery is more heavily weighted with items that re-
present the Skill measure, so that a sum for the total battery would favor programs
which emphasize teaching skills. In contrast, the Abstract measures which would
seem to be closer to the objectives of the inquiry-oriented programs are repre-
sented by a small number of items which would easily be overwhelmed in the total

battery.

Beyond the issue of program evaluation, when these measures differing in
nature are used to examine the growth of pupil subgroups differing in ethnic
group and socio-economic status, the minor differences in growth in either
Concrete or Abstract measures are so inconsistent as to lead one to conclude
that there is no real difference. Subgroups start at different levels and
finish at different levels, but grow at very similar rates for both Abstract
and Concrete pupil measures. This seems a very hopeful finding for a compen-
satory education program.

Relations Between Observational Data and Pupil Growth

In contrast to the use of observational data in program discrimination, the
relations of the observational measures to measures of pupil growth are scattered
and often inconsistent. Several problems contribute to difficulty in drawing
dependable conclusions. Growth is the problem being studied, but growth measures
are much less reliable, and correlate much less strongly with each other and with
other measures; than scores which represent standing at a point in time. Thorndike
(1966) estimates that the correlation of a child's standing at the beginning of
the year with his growth during the year is probably no more than +.10. In con-
trast, correlations between standings for elementary pupils commonly approach +.70.
In these data, growth measures have been pooled, classroom by classroom, so that
reliabilities have been increased, but to judge from other studies, high correlations
are not to be expected.

The small number of classrooms in each group presented another problem -- the
largest number of first grades is 20, which requires a correlation of .44 to be
significant. Numbers range downward to 9 which requires a correlation of .67 to
be significant. It seems likely, then, that significant correlations will be found
only for real relationships which are augmented by chance variation which happens

to fall in the direction of increasing the correlation. Many real relationships, then, will be obscured by the overlay of chance correlations associated with small samples. (Kindergartens have a slightly larger number of classrooms, but only rarely correlate significantly with any measure, the number of significant correlations being reasonably attributable to chance.)

Against this backdrop of uncertainty, a few relationships will be cited which are identified by several significant correlations, supported by large but nonsignificant ones.

Complex-Abstract growth appears to be fostered by several measures which theoretically complement each other. Experimental Teaching in the Deweyan sense appears to be such a measure -- pupils are confronted by difficult problems, encouraged to hypothesize and examine alternatives, but required to evaluate·the alternatives they have suggested (TPOR 2). Another is Intrinsic Motivation (TPOR 5), reflecting the involvement of the pupil in the ideas and the activities themselves, rather than being motivated by teacher direction and evaluation. Additional support seems to come from teaching which is not narrowly focused on subject matter but is involved with the broader experience of the child and perhaps with his world out of school (CogTax 3). The measure which seems to be most strongly supportive of Abstract growth is so brief as to be enigmatic, but appears to reflect effective personal communication between teacher and pupils in which pupils appear to feel the teacher's involvement with them (FLACCS 7). In contrast, one measure appears to inhibit Complex-Abstract growth: a measure in which the teacher occasionally exercises very close control over pupils in an otherwise little structured setting (FLAC 6). It is not clear, however, whether it is the close control or the inconsistent teacher behavior, or both, which have their effect. All in all, this seems a reasonably coherent pattern of classroom behaviors, and one which it seems reasonable to find related to Abstract growth.

Pupil growth in Skill appears to be related to a number of the same classroom behaviors cited for Abstract growth. Intrinsic motivation (TPOR 5), Effective Personal Communication (FLACCS 7), and Moderately Focused (partly non-school related) Teaching (CogTax 3) all appear to support the growth of skill. However, it is not supported by the Experimental Teaching (TPOR 2) which supported Abstract growth.

Fewer measures appear to relate to Simple-Concrete growth. Teaching which is very specifically focused on lower level academic skills promotes Concrete growth (CogTax 1), but more broadly focused teaching (CogTax 3) appears to inhibit it. Effective Personal Communication (FLACCS 7) supports it, but extreme freedom of action of pupils in the classroom (FLAC 5) also appears to inhibit it.

There appears to be a kind of general trend which runs through these relationships which suggests that pupil Abstract growth tends to be supported by less controlled, less focused classroom activities than support Skill growth, and that Concrete growth is supported by still more specific teacher activities. Given the small samples, and the inconsistencies of the data, however, this is to a degree tentative.

A somewhat surprising result is the failure of variables reflecting negative emotional climate to relate to measures of pupil growth. In contrast, two factors which reflect positive involvement of the teacher with pupils related to growth; one was a very powerful correlate and the other approached significance. A possible explanation for the difference in pupil response may be that disadvantaged children live with negative affect so commonly that it ceases to have much impact. If so, it would be reasonable to find positive affect variables having impact, as a kind of contrast effect.

In looking at the measures as a group and their relations with pupil growth, it is interesting to note that factors which discriminate programs most strongly,

and which represent major portions of the variance in the observational data, are not ones which relate strongly to pupil growth. The extent to which these factors discriminate programs implies that they represent an important dimension along which programs differ -- perhaps a dimension whose poles are represented in contemporary educational thought by the advocacy of programmed instruction on the one hand, and "open" classrooms on the other.

A problem in the analyses of these data which may explain the lack of relation just cited as well as the lack of significant results in general, may be the tendency for these relationships to be nonlinear. An examination of the first year data suggests that measures of teacher control and pupil freedom relate in nonlinear fashion to pupil growth. The numbers of classrooms within any grade level group were too small to permit a reasonable test of the hypothesis, but several bits of data point to this possibility.

In summary, observation methods appear to be significant discriminators of at least some program objectives, but the relations between observational data and pupil growth are less clear perhaps as a consequence of smaller samples of pupil data and the possibility of nonlinear relations. The Complex-Abstract, Skill and Simple-Concrete subgroups of pupil growth measures which were created appeared to relate to program differences and observational measures in meaningful ways.

References

Amidon, E. J., & Hough, J. B. Interaction analysis: theory, research and application. Reading, Mass.: Addison-Wesley, 1967.

Bane, R. K. Relationships between measures of experimental, cognitive and affective teaching behavior and selected teacher characteristics. Unpublished doctoral dissertation, Univ. of Fla., Gainesville, 1969.

Bereiter, C. Some persisting dilemmas in the measurement of change. In Harris, C. W. (Ed.), Problems in measuring change. Madison: University of Wisconsin Press, 1963.

Bloom, B. S., et al (Ed.) Taxonomy of educational objectives Handbook I: cognitive domain. New York: McKay, 1956.

Brown, B. B. The experimental mind in education. New York: Harper & Row, 1968.

Dixon, W. J. BMD biomedical computer programs (2nd Ed). Los Angeles: Health Sciences Computing Facility, Univ. of Calif., 1970.

Flanders, N. A. Teacher influence, pupil attitudes, and achievement. Coop. Res. Monograph No. 12, OE-25040, US Dept. of HEW, 1965.

Flanders, N. A. Analyzing teacher behavior. Reading, Mass.: Addison-Wesley, 1970.

Fowler, Beverly D. Relation of teacher personality characteristics and attitudes to teacher-pupil rapport and emotional climate in the elementary classroom. Unpublished doctoral dissertation, Univ. of S. Carolina, 1962.

Glass, G. V., and Maguire, T. O. Abuses of factor scores. Am. Ed. Res. Jour., 1966, 3, 297-304.

Hayes, D. P., & Grether, Judith. The school year and vacations: When do students learn? Paper presented at the Eastern Sociological Assoc. Con., New York City, April, 1969.

Honeycutt, Joan K. Relationship between disadvantaged pupil achievement and Reciprocal Category System-assessed reward and punishment. Unpublished doctoral dissertation, Univ. of Fla., Gainesville, 1971.

Horn, J. L. An empirical comparison of various methods for computing factor scores. Ed. Psychol. Meas., 1965, 25, 313-322.

Jennrich, R. I., & Sampson, P. F. Rotation for simple loadings. Psychometrika, 1966, 31, 312-323.

Katz, L. G., Peters, D. L., & Stein, N. S. Observing behavior in kindergarten and preschool classes. Childhood Education, 1968, 44, 6, 400-405.

Masling, Joseph & Stern, George. Effect of the observer in the classroom.
    Jour. of Ed. Psychol., 1969, 60, 351-354.

Medley, D. M., & Mitzel, H. E. Measuring classroom behavior by systematic
    observation. In Gage, N. L. (Ed.), Handbook of research on teaching.
    Chicago: Rand McNally & Co., 1963, 247-328.

Ober, R. L., et al. The development of a reciprocal category system for
    assessing teacher-student classroom verbal interaction. Paper presented
    at the American Educational Research Association, Chicago, 1968.

Samph, T. Observer effects on teacher behavior. Unpublished doctoral dis-
    sertation, Univ. of Michigan, Ann Arbor, 1968.

Sanders, N. M. Classroom questions: What kind? New York: 'Harper &
    Row, 1966.

Sears, R., Rau, L., & Alpert, R. Identification and child rearing. Stanford:
    Stanford Univ. Press, 1964.

Soar, R. S. An integrative approach to classroom learning. NIMH project'
    numbers 5-R11 MH 01096 to the Univ. of South Carolina, and 7-R11 MH 02045
    to Temple Univ., Philadelphia, Pa., 1966. Eric document ED 033 749.

Soar, R. S. Optimum teacher-pupil interaction for pupil growth. Educ.
    Leadership Res. Supplement, 1968, 1, 275-280.

Thorndike, R. L. Intellectual status and intellectual growth. Jour. of Ed.
    Psychol., 1966, 57, 121-127.

Wood, S. E. An analysis of three systems for observing classroom behavior.
    Unpublished doctoral dissertation, Univ. of Fla., Gainesville, 1969.

Appendix A

Equipment for Classroom Tape Recording

The final procedure involved the use of a moderately priced tape recorder and a Cardiod microphone (so-called because a graph of its sensitivity is somewhat heart-shaped). This class of microphones is broadly sensitive to the front, but sensitivity declines sharply to the rear, with the instrument almost completely insensitive directly to the rear. In use the microphone was not so much directed at the sound to be recorded, but away from competing sounds. In addition, these professional microphones are materially more sensitive than those furnished with tape recorders.

A further increase in recording quality was gained from the use of "high-output" tape which is more sensitive to faint signals.

The choice of batteries for the tape recorder also was an important factor. Full voltage is necessary to obtain the best possible recording. Zinc-acid batteries (ordinary flashlight batteries) begin to decline in voltage after a few minutes' use and decline steadily. GE rechargeable batteries produce only 1.3 volts instead of 1.5 at full charge, so that a set of five batteries in series produces a voltage a full volt below nominal value. Mallory batteries and chargers were selected because they produce full voltage which is sustained for extended periods of time.

A final change was the provision of moderate quality earphones. Stereo earphones, rewired to function monaurally, were found to increase intelligibility over use of the speaker in the recorder, or earphones intended for transcription.

The result of the various changes was that, in general, anything that a live observer could hear and understand in the classroom became codeable from tape.

Appendix B

Statistical Procedures Used in Calculating Pupil Gain Measures

Status vs Gain

In assessing pupil growth, it seemed important to use a measure of gain rather than a measure of standing at the end of the year. Gain measures are more sensitive reflections of change than status measures, since the individual is being compared with himself and his earlier status. Raw gain (the difference between post-test and pre-test) however, is generally an undesirable way to reflect change, because of regression effect. The concept of regression effect is that if two measures are less than totally reliable, a person who scores low on the first test will tend to score somewhat higher on the second; and a person who scores high on the first test will tend to score lower on the second, assuming no real change from one test to the other. A parallel statement is that a spurious negative correlation can be expected between the pre-test score and the difference from pre- to post-score. When a period of time intervenes between pre- and post-test and growth occurs, the effect is less obvious but the spurious negative correlation can be expected to remain. Its effect will be to increase the apparent growth of initially low scoring pupils, and to decrease the apparent growth of initially high scoring pupils. As a matter of fact, in the first grade sample of data analyzed the first year, the median correlation between pre-test score and raw gain was -.61 across the 13 subtests of the Metropolitan Readiness Test, the Early Childhood Inventory, and the Caldwell Soule. Since this negative correlation is a measurement artifact and clouds relations of these measures with others, it seems desirable to remove it.

Regressed Gain

Regressed gain is a way of adjusting out regression effect. Its effect

is to adjust out that portion of gain which is associated with pre-test score, thus eliminating the spurious relationship. In the course of this adjustment, the gains of initially low-scoring pupils are decreased, and the gains of initially high-scoring pupils are increased. This adjustment, however, raises other difficulties in turn. A regression effect takes place toward the mean of the population of which a given pupil is a member. The further away he is from the mean of that population, the greater is the adjustment which is made. The problem, then, is that if members of two populations which differ in mean score are combined in an overall analysis, the gains of the members of the lower standing population will be reduced more than they should be, and the gains of the members of the high-scoring population will be increased more than they should be. The usual solution to this problem, of course, is to carry out a separate analysis for each population.

In the first year's data, this could not be done because demographic data were available on too few pupils, so regressed gain was calculated separately for the total kindergarten group and the total first grade group. Within the second year data, four subpopulations appeared to warrant separate analysis, identified by ethnic group (white and black), and by social status (defined by the report of whether the pupil was entitled to full Follow Through services). The difficulty, however, was that a different test battery had been administered to kindergarten pupils, and entering first grade pupils (those without previous Follow Through kindergarten experience), from that administered to nonentering first grade pupils (those with previous kindergarten experience). When these three subgroups were each divided into four subgroups, the numbers of cases became too small for separate analysis -- as small as an N of three in one case.

A different alternative seemed reasonable if mean gain for each subgroup

appeared to be similar to that of the other subgroups within a grade level group-

ing. If mean gain from pre- to post- is similar for two groups, and if variabili-

ties are similar, then the regression of post-score on pre-score may reasonably

be assumed to be the same, and the groups can be combined if the differences in

initial standing can be eliminated. This was the approach that was followed.

Inspection of the pre-post means across the various measures did not show consis-

tent differences from subgroup to subgroup in amount of gain (see Table 4).

Differences in the levels of the regression lines were adjusted out for each item

or subtest in the following way. If, for example, the low social status white

group stood three points lower than the high social status white group in pre-test

mean, then three points was added to each pre- and post-test score in the white low

social status group. In this way, the amount of change from pre- to post- was

left unchanged, but the level of the regression line was shifted. Similar adjust-

ments were made on all groups, so that the four regression lines should, in effect,

become a common one. Regressed gain was then calculated for the combined group,

within each grade level group.

Factor Analysis of Gain Scores vs Status Scores

Analysis of the subtest scores from the first year's data indicated that

the factor structure of regressed gain scores was appreciably different from that

of either pre- or post-scores. It was only in the analysis of regressed gain

scores that the simple-complex structure emerged clearly. If one is interested

in gain, he should analyze gain. As Bereiter (1963) has indicated, items which

are selected to measure standing at some point in time are likely to be items

which are quite stable, and consequently not good measures of change. It seems

possible, then, that the factor analyses of regressed gain measures in the first

year's data may have been identifying measures which were more sensitive to change

than those which failed to load.

These several considerations led to the assumption that the analysis of items and subtests to create new pools in the second year's data paralleling the first year factors should be done using measures of change. The difficulty with this procedure, however, was the cumulative loss of reliability from change measures being much less reliable than the status measures from which they are derived, and items being less reliable than subtests.

In one sense, the use of the term "items" is inappropriate in that a number of items on the WRAT (for instance, Word Reading, N = 20) have as long or longer scales than subtests of the Metropolitan (Word Meaning, N = 7). But there were items with one and two point scales.

Before items were factor analyzed, those which "topped out" or had very low variability were eliminated. Despite the problems of measures with varying scale length occurring in the same analysis and the reliability problems cited earlier, reasonably clear structures were obtained from factor analyses of the kindergarten and entering first grade, analyzed separately. In addition to the simple and complex factors found the first year, a third factor emerged which was unexpected. It appeared to represent skills-learning such as reading, spelling, and arithmetic (the three R's!). It was apparently associated with the addition of the Wide Range Achievement Test (WRAT), which is heavily weighted with these skills. Incomplete factor scores from these analyses were used as the measures of pupil regressed gain.

For the nonentering first grade data, items and subtests which appeared to be abstract came together with a series of low-variance items which appeared to measure simpler kinds of learning. Inspection of the data also suggested a tendency for items to be grouped on the basis of whether they were individually or group administered. Various combinations of items and subtests were combined and factor analyzed, but no really satisfactory structure emerged. Instead, on the basis of

the factor analysis and also a priori judgment, composites of items and subtests

were formed to represent group and individually administered, simple-concrete,

skill, and complex-abstract measures. An item analysis of all items against

these composites was carried out, and items added to the composite accordingly.

Appendix C

Table C-1

Pupil Data Status - 1970 Sample

	Kindergarten	First Grade Entering	Nonentering	Total
Classrooms Observed	30	15	25	70
Classrooms Missing Pre or Post Data	10	6	13	29
Classrooms Missing Demographic Data	1	0	0	1
Classrooms Used in Pupil Analyses	19	9	12	40

Appendix D

Appendix D

Program Descriptive Data

Pupil Growth Data

Pupil growth data are not reported for the programs observed the first year. Demographic data were available for only a small fraction of the pupils on whom test data were obtained, and the estimation of regressed gain was carried out for the total group. Since this calculation for different groups combined creates a bias which is different for higher and lower standing sub-groups (see Appendix B), and since programs differ in the ethnic compositions of the pupils served, comparisons would not have been meaningful.

In the 1970 data, demographic data were available on most pupils for whom test scores were obtained, so that regressed gain could be estimated with the effect of economic status and ethnic group adjusted out. Although program comparisons are reported for the 1970 data, the numbers of classrooms involved are too small for meaningful comparisons of programs. The data are reported only for the light shed on the sensitivity of the different pupil gain measures to different program objectives. The changes in standing for the same groups of pupils on the different measures suggest that different aspects of pupil growth are being measured. The means reported are deviations around the grand mean of regressed gain for all pupils.

The data for kindergarten growth scores are shown in Table D-1. The BE program stands high for Complex-Abstract growth, stands high along with the Nimnicht program for Skill growth, but the Comparison group stands high for Simple-Concrete growth.

In the entering first grade data (Table D-2), the only significant

Table D-1

Multiple Range Tests - Kindergarten Growth Scores (1970)

Complex-Abstract					Skill				
Sponsor	Mean	NSR*	S.D.	N[1]	Sponsor	Mean	NSR*	S.D.	N
BE	50.70		.07	3	BE	50.51		.09	3
Nimnicht	50.07		.31	4	Nimnicht	50.21		.32	4
Comp.	49.91		.76	3	Comp.	49.82		.16	3
EDC	49.88		.13	2	Gotkin	49.80		.10	4
Gotkin	49.80		.22	4	Bank St.	49.79		.06	2
Bank St.	49.42		.16	2	EDC	49.69		.01	2

$F = 3.46$ N.S. $F = 8.81$ N.S.

Simple-Concrete				
Sponsor	Mean	NSR*	S.D.	N
Comp.	50.77		.31	3
EDC	50.38		.34	2
Nimnicht	50.24		.61	4
BE	50.11		.35	3
Gotkin	49.51		.61	4
Bank St.	49.47		.51	2

$F = 2.98$ N.S.

*Nonsignificant ranges.

[1]N is the number of classrooms. One program with only one classroom is omitted.

Table D-2

Multiple Range Tests - Entering First Grade Growth Scores (1970)

	Complex-Abstract					Skill			
Sponsor	Mean	NSR*	S.D.	N[1]	Sponsor	Mean	NSR*	S.D.	N
Bank St.	50.14		.15	3	Bank St.	50.25		.21	3
Comp.	49.95		.01	2	EDC	50.06		.31	2
EDC	49.93		.02	2	Comp.	49.79		.35	2
Tucson	49.88		.18	2	Tucson	49.68		.41	2

F = 2.30 N.S.   F = 1.71 N.S.

	Simple-Concrete			
Sponsor	Mean	NSR*	S.D.	N
Tucson	50.33		.16	2
Comp.	50.12		.34	2
Bank St.	49.96		.08	3
EDC	49.76		.04	2

F = 3.95 N.S.

*Nonsignificant ranges.

[1]N is the number of classrooms.

separation is for Simple-Concrete measure in which Tucson pupils show significantly more growth than EDC pupils. Program orders change from measure to measure, but no other differences are significant.

In the nonentering first grade data (Tables D-3, D-4, and D-5), the measures are broken into Group Administered, Individually Administered, and Total, in addition to the levels of abstraction which were used in the first two groups. The BE program stands highest on all three of the skill scores, in each instance differing significantly from the lowest standing program. But on the Group Complex-Abstract measure, the BE program stands lowest, and significantly different from the PE program which stands highest. Again, the only conclusions which are appropriate have to do with the differential characteristics of the measures, not the relative standings of programs.

The data do appear to point to the usefulness of measures such as these for representing differences in program objectives in program evaluation.

Table D-3

Multiple Range Test - Nonentering First Grade Growth Scores (1970)

Group Administered

Complex-Abstract					Skill				
Sponsor	Mean	NSR*	S.D.	N[1]	Sponsor	Mean	NSR*	S.D.	N
PE	50.72		0.32	3	BE	50.35		0.91	2
Nimnicht	50.08		0.06	2	Nimnicht	50.26		0.88	2
Tucson	50.06		0.33	3	PE	50.22		1.63	3
BE	49.34		1.20	2	Tucson	49.67		1.03	3

$F = 2.48$ N.S. $\qquad$ $F = 4.27$ N.S.

Simple-Concrete				
Sponsor	Mean	NSR*	S.D.	N
PE	50.27		0.33	3
BE	50.10		0.87	2
Nimnicht	50.03		1.29	2
Tucson	49.82		1.02	3

$F = 1.15$ N.S.

*Nonsignificant ranges.

[1]N is the number of classrooms. Two programs with one classroom each are omitted.

Table D-4

Multiple Range Test - Nonentering First Grade Growth Scores (1970)

Individually Administered

Complex-Abstract					Skill				
Sponsor	Mean	NSR*	S.D.	N[1]	Sponsor	Mean	NSR*	S.D.	N
PE	.50.24		1.76	3	BE	50.66		4.90	2
Nimnicht	50.13		0.25	2	Nimnicht	50.65		0.35	2
BE	50.00		1.09	2	PE	50.26		7.65	3
Tucson	49.83		0.27	3	Tucson	49.07		2.43	3

$F = 1.75$ N.S. $F = 4.71$ N.S.

Simple-Concrete				
Sponsor	Mean	NSR*	S.D.	N
BE	50.46		2.16	2
PE	50.39		0.40	3
Nimnicht	50.01		2.59	2
Tucson	49.72		1.14	3

$F = 1.19$ N.S.

*Nonsignificant ranges.

[1]N is the number of classrooms. Two programs with one classroom each are omitted.

147

Table D-5

Multiple Range Test - Nonentering First Grade Growth Scores (1970)

Total

Complex-Abstract					Skill				
Sponsor	Mean	NSR*	S.D.	N[1]	Sponsor	Mean	NSR*	S.D.	N
PE	50.32		2.08	3	BE	50.56		5.81	2
Nimnicht	50.12		0.32	2	Nimnicht	50.51		0.53	2
BE	49.89		2.29	2	PE	50.24		9.28	3
Tucson	49.87		0.14	3	Tucson	49.28		3.46	3

F = 1.92 N.S.   F = 4.73 N.S.

Simple-Concrete				
Sponsor	Mean	NSR*	S.D.	N
PE	50.33		0.16	3
BE	50.28		3.03	2
Nimnicht	50.02		3.88	2
Tucson	49.77		2.16	3

F = 1.18 N. S.

*Nonsignificant ranges.

[1]N is the number of classrooms. Two programs with one classroom each are omitted.

Classroom Observation Data

In comparing classroom observation data from the first year to the second, several considerations seem important. First, more than half of the teachers in the sample the first year were replaced the second year. Replacement ranged from none for one program to complete replacement for another (see Table 2). A number of teachers changed their teaching style in major ways from the first to the second year, as seems reasonable to expect as programs mature. In addition, evidence of cross-fertilization and drawing together of programs seemed clear on occasion -- the presence of Gotkin materials in other program classrooms, for example. But, in contrast, in some urban areas, local problems contributed to tensions in the school which were reflected in the classroom.

Despite these various influences, a considerable degree of similarity can be seen in the order of programs from one year to the next.

Teacher Practices Observation Record

Factor 1 - Pupil Free Choice vs Teacher Structured Activity - In the first year's data, EDC classrooms were separated from all others as showing greater amounts of pupil choice of activity; in the second year's data, Nimnicht classrooms joined EDC classrooms in separating from other programs (see Table D-6). Freedom of choice appears to be central to the EDC program as it is described, and these data appear to agree. Probably the high standing of the Nimnicht program relates to the fact that pupils spend a considerable portion of the day working with various kinds of audio-visual equipment of their own choice.

In both years, Becker-Engelmann classrooms are identified as those in which most teacher structuring takes place. The emphasis on skill development through the use of programmed materials would seem to require that BE classrooms be rather highly structured with little pupil freedom of choice.

Table D-6

Multiple Range Test - Teacher Practices Observation Record

Factor 1 - Pupil Free Choice vs Teacher Structured Activity

	1969					1970			
Sponsor	Mean	NSR*	S.D.	N	Sponsor	Mean	NSR*	S.D.	N
EDC	59.56		5.41	8	Nimnicht	58.49		5.45	8
Nimnicht	53.36		3.49	8	EDC	57.30		6.61	8
Bank St.	51.29		6.77	8	Tucson	51.90		4.20	8
Gotkin	51.02		3.60	8	Bank St.	50.01		5.83	8
Tucson	50.53		6.20	8	PE	48.75		6.06	8
PE	49.96		3.92	8	Gotkin	48.66		4.60	8
Comp.	45.95		5.15	14	Comp.	46.10		4.38	14
BE	41.38		3.38	8	BE	43.65		3.34	8

$F = 9.83$ $p < .01$ $F = 8.76$ $p < .01$

*Nonsignificant ranges.

Table D-7

Multiple Range Test - Teacher Practices Observation Record

Factor 2 - Experimental Teaching

	1969					1970			
Sponsor	Mean	NSR*	S.D.	N	Sponsor	Mean	NSR*	S.D.	N
EDC	54.15		2.48	8	Nimnicht	55.05		5.43	8
Bank St.	51.29		4.63	8	PE	52.84		3.97	8
PE	51.05		4.08	8	Tucson	51.79		7.77	8
Tucson	50.95		3.76	8	EDC	51.38		6.03	8
Nimnicht	49.67		1.96	8	Bank St.	50.38		6.59	8
Gotkin	49.43		2.32	8	Comp.	48.86		5.36	14
BE	48.02		1.78	8	Gotkin	48.78		4.08	8
Comp.	47.39		3.10	14	BE	45.28		2.45	8

$F = 4.25$ $p < .01$ $F = 2.46$ $p < .05$

*Nonsignificant ranges.

This factor appears to be a major dimension along which programs differ, but it does not relate to pupil growth.

Factor 2 - Experimental Teaching - The order of programs differed somewhat from the first to the second year's data (Table D-7). In the first year's data, the EDC program stood highest on the factor with Bank Street, PE, and Tucson, with Comparison, Becker-Engelmann, Gotkin and Nimnicht classrooms in a lower grouping. In the second year's data, the Nimnicht program moved from the lower group to the upper group, and Bank Street became the program common to both groups. The shift of the Nimnicht program may partially reflect the replacement of seven of eight teachers from the first to the second year's sample. Teacher variability within program is high in comparison to differences between means.

Factor 3 - Less vs More Differentiation - With the exception of the Becker-Engelmann program, programs go in roughly the same order in both years' data (see Table D-8). Ignoring it for the moment, Comparison groups are at the end of the scale showing less differentiation of activities, with EDC, Tucson, and Nimnicht at the other end of the scale. Becker-Engelmann classrooms reversed position from least differentiation in the first year's data to most differentiation in the second year's data. This change took place as a consequence of an observer ground rule change. In the first year's data, observers tallied only what they could see or hear, but since subgroups sometimes went to different rooms (where only this subgroup was present), such a group was tallied as undifferentiated. In the second year's data, however, a subgroup was tallied as differentiated if the observers knew that other subgroups from that classroom were engaged in other activities, regardless of the fact that these behaviors were not observable.

Factor 4 - Nonexperimental Teaching - This factor discriminated between

Table D-8

Multiple Range Test - Teacher Practices Observation Record

Factor 3 - Less vs More Differentiation

	1969					1970			
Sponsor	Mean	NSR*	S.D.	N	Sponsor	Mean	NSR*	S.D.	N
BE	55.45		5.30	8	Comp.	55.97		7.11	14
Comp.	53.90		5.50	14	Gotkin	54.60		5.66	8
Gotkin	51.95		3.54	8	Bank St.	49.33		6.39	8
Bank St.	51.75		4.75	8	PE	49.00		4.84	8
PE	49.38		3.04	8	Tucson	47.83		6.55	8
EDC	45.93		6.20	8	Nimnicht	47.63		4.13	8
Nimnicht	45.08		5.34	8	BE	47.13		5.19	8
Tucson	44.65		4.88	8	EDC	45.95		3.23	8

$$F = 5.67 < p \ .01 \qquad\qquad F = 4.21 \quad p < .01$$

*Nonsignificant ranges.

Table D-9

Multiple Range Test - Teacher Practices Observation Record

Factor 4 - Nonexperimental Teaching

	1969					1970			
Sponsor	Mean	NSR*	S.D.	N	Sponsor	Mean	NSR*	S.D.	M
BE	58.29		3.70	8	BE	57.28		4.71	8
Comp.	51.51		4.80	14	Comp.	53.98		5.89	14
Bank St.	49.78		2.78	8	Bank St.	53.75		6.65	8
PE	49.69		2.89	8	Gotkin	50.70		4.07	8
Nimnicht	47.98		3.83	8	Tucson	48.24		3.32	8
Gotkin	47.37		2.93	8	PE	46.65		2.81	8
EDC	47.22		2.27	8	EDC	45.53		3.96	8
Tucson	47.02		2.07	8	Nimnicht	44.03		3.13	8

$$F = 9.66 \quad p < .01 \qquad\qquad F = 8.67 \quad p < .01$$

*Nonsignificant ranges.

programs significantly in both year's data (Table D-9). Becker-Engelmann class-rooms stand high, alone in the first year's data, and together with Comparison and Bank Street classrooms in the second. The high standing for BE seems to represent the objectives of the program. The lower end of the factor shows con-siderable variability in order, except for EDC which retains a position toward the lower end both years. This seems to agree with objectives as well.

Factor 5 - Intrinsic Motivation in a Task Setting - In the first year's data, EDC and Bank Street were separated off at the upper end; the second year, BE, Bank Street, and Gotkin were at the upper end of the scale (Table D-10). Although contingency management programs make extensive use of extrinsic motiva-tion, the sequence of activities which build involvement have warranted checking the item for intrinsic motivation with some frequency in these classrooms.

Factor 6 - Positively Focused Teaching - Programs are widely separated by this factor and the F's are the largest for both years (Table D-11). The BE program stands high and is widely separated from the next highest program both years. The EDC program stands lowest and is widely separated from the rest of the programs in the first year's data, but is joined by the Nimnicht and Tucson programs in the second year's data. These findings appear to agree with the ob-jectives of the programs. The teacher direction present in a teacher operated learning program sets off BE classrooms. In contrast, EDC, Nimnicht, and Tucson classrooms appear there as ones in which the teacher does not take a central role in "teaching" children, but encourages pupil independence in the learning process.

Table D-10

Multiple Range Test - Teacher Practices Observation Record

Factor 5 - Intrinsic Motivation in a Task Setting

Sponsor	1969 Mean	NSR*	S.D.	N	Sponsor	1970 Mean	NSR*	S.D.	N
EDC	52.18		3.93	8	BE	53.35		6.51	8
Bank St.	51.81		3.20	8	Bank St.	52.70		5.76	8
BE	50.66		4.90	8	Gotkin	52.70		5.24	8
Gotkin	50.58		3.30	8	Nimnicht	50.90		4.90	8
PE	50.58		2.14	8	Tucson	49.88		8.17	8
Nimnicht	49.91		2.52	8	EDC	49.70		7.33	8
Comp.	48.45		2.64	14	PE	49.65		4.35	8
Tucson	47.06		3.84	8	Comp.	46.57		4.53	14

F = 2.24 p < .05   F = 1.54   N.S.

*Nonsignificant ranges..

Table D-11

Multiple Range Test - Teacher Practices Observation Record

Factor 6 - Positively Focused Teaching

Sponsor	1969 Mean	NSR*	S.D.	N	Sponsor	1970 Mean	NSR*	S.D.	N
BE	60.88		5.22	8	BE	60.99		2.56	8
Comp.	52.12		3.97	14	Gotkin	52.71		4.03	8
PE	50.93		2.46	8	Comp.	50.97		4.92	14
Bank St.	49.08		3.22	8	Bank St.	50.54		5.27	8
Gotkin	48.57		2.46	8	PE	49.59		4.82	8
Nimnicht	47.58		3.80	8	Tucson	45.78		3.12	8
Tucson	46.71		3.42	8	Nimnicht	45.39		5.21	8
EDC	42.53		3.29	8	EDC	44.31		4.27	8

F = 17.69 p < .01   F = 11.74 p < .01

*Nonsignificant ranges.

154

Florida Affective Categories System (1969)

Factor 1 - Free Choice vs Structured Learning in Groups - The finding
that EDC classrooms stand high in freedom of choice for pupils seems closely
aligned with program rationale, as does the standing of the BE program at the
low end of the factor (Table D-12). The standing of Comparison classrooms
toward the lower end of the factor agrees with the common conception of class-
rooms in general. However, the nature of the structuring done by the Compari-
son teacher is quite different from that in BE classrooms; subjectively this is
true, and other factors support and clarify this.

Factor 2 - Intermittent Work vs Group Singing and Games - The multiple
range test (Table D-12) showed two homogeneous groups, indicating that the Nim-
nicht program differs from the others at the high end, perhaps reflecting the
transition of activities as pupils move from one to another of the various
learning activities. Comparison classrooms are set off at the lower end of the
factor, perhaps reflecting a larger proportion of activities of the total class.
The factor does not seem particularly clear, and has one of the lowest F's of
the FLAC factors.

Factor 3 - Warm Emotional Climate - The data of Table D-13 show two
homogeneous subsets which indicate that Comparison classrooms differ from Nim-
nicht classrooms which stand highest on the factor, but program classrooms do
not differ one from another. Again, the overall F was not significant. This
is an objective which all programs value.

Factor 4 - Teacher Negative Affect vs Pupil Work and Socialization with
Adults - The data of Table D-13 show two homogeneous subsets, indicating that
the Tucson, BE, and EDC programs show less teacher negative affect and/or more
pupil work and socialization than do Comparison classrooms. The avoidance of

Table D-12

Multiple Range Test - Florida Affective Categories - 1969

Factor 1 - Free Choice vs Structured Learning in Groups

Factor 2 - Intermittent Work vs Group Singing or Games

Sponsor	Mean	NSR*	S.D.	N	Sponsor	Mean	NSR*	S.D.	N
EDC	57.61		6.47	8	Nimnicht	53.89		5.31	8
Nimnicht	55.76		4.25	8	Bank St.	50.54		3.03	8
PE	53.11		1.83	8	Tucson	50.31		5.81	8
Bank St.	52.74		5.76	8	PE	50.25		6.67	8
Gotkin	52.70		2.39	8	EDC	49.81		6.17	8
Tucson	51.63		2.72	8	BE	49.28		6.84	8
Comp.	47.36		5.09	14	Gotkin	47.19		3.56	8
BE	45.51		3.38	8	Comp.	46.79		8.06	14

F = 7.60**

F = 1.21

Table D-13

Multiple Range Test - Florida Affective Categories - 1969

Factor 3 - Warm Emotional Climate

Factor 4 - Teacher Negative Affect vs Pupil Work and Socialization with Adults

Sponsor	Mean	NSR*	S.D.	N	Sponsor	Mean	NSR*	S.D.	N
Nimnicht	54.20		4.65	8	Comp.	54.11		5.73	14
EDC	53.06		5.34	8	Nimnicht	52.44		8.03	8
BE	52.38		6.99	8	PE	52.18		4.32	8
Tucson	52.11		3.60	8	Gotkin	50.04		3.47	8
PE	51.03		5.57	8	Bank St.	49.70		6.69	8
Bank St.	50.45		6.01	8	Tucson	47.70		4.74	8
Gotkin	49.48		3.25	8	EDC	47.63		3.29	8
Comp.	46.93		8.47	14	BE	47.10		6.45	8

F = 1.54

F = 2.13

correction or criticism of pupils in the Tucson program is an important aspect of program rationale; and the BE program stresses reinforcement of desired be-haviors so that criticism of pupils is relatively infrequent. A supportive climate for pupils is also an important aspect of the EDC program.

Factor 5 - Pupil Free to Withdraw, Seek Reassurance, Pretend, or Express Negative Affect. The data of Table D-14 show that the EDC program is widely separated from other programs and Comparison classrooms. The data illustrate the freedom and autonomy of the pupils in the EDC program.

Factor 6 - Teacher Neutral Control vs Teacher Support in Task Settings - The results of the multiple range test show three homogeneous groupings, and the overall F is significant (see Table D-14). At first it seems paradoxical to find the EDC program, in which there is a great deal of freedom for pupils, at the top of this factor, and the BE program, in which there is little autonomy for pupils, at the bottom of the factor. The issue seems to be the kind of freedom. As indicated in the body of the report, Factor 1 represents a differ-ent dimension, interpreted as a lack of structure in which, in extreme cases, the teacher apparently feels occasional need to intervene firmly (reflected in this factor). The structure of the BE program apparently minimizes the need for control as this factor identifies it.

Table D-14

Multiple Range Test - Florida Affective Categories - 1969

Factor 5 - Pupil Free to Withdraw Seek
Reassurance, Pretend or
Express Negative Affect

Factor 6 - Teacher Neutral Control vs Teacher
Support in Task Settings

Sponsor	Mean	NSR*	S.D.	N	Sponsor	Mean	NSR*	S.D.	N
EDC	58.26		4.77	8	EDC	55.21		3.26	8
Bank St.	52.88		7.53	8	Gotkin	53.45		6.34	8
PE	52.21		4.51	8	Comp.	52.13		6.37	14
Nimnicht	52.13		3.95	8	Tucson	49.98		3.95	8
Gotkin	51.48		3.09	8	Bank St.	49.80		4.26	8
Tucson	50.63		4.02	8	PE	48.83		6.95	8
Comp.	50.21		5.60	14	Nimnicht	46.98		5.00	8
BE	48.53		4.78	8	BE	46.85		6.62	8

$F = 2.72*$ $F = 2.37*$

158

Florida Climate and Control System (1970)

Factor 1 - Teacher Negative Control - This factor separates programs into two groups as shown in Table D-15. The Nimnicht and Tucson programs are separated from other programs at the lower end of the scale, and the Gotkin and Bank Street programs are separated off at the upper end. It is the subjective impression of observers that the occurrence of this sort of classroom in the latter two programs appeared primarily in one school for each program -- schools which were sharply different the third year of the program. In a sense, it seems unfortunate to report these data in relation to programs in that they appear to be largely school effect. Both were schools undergoing change, and this stress was apparently reflected in classrooms. In addition, the Gotkin program was the only one for which the entire sample was from large cities, where stresses more often occur.

Factor 2 - Pupil Freedom vs Teacher Structured Activity - The pattern of separations for this factor (Table D-15) is quite similar to that of FLAC Factor 1 for the 1969 data. The Nimnicht and EDC programs are set off at the top of the factor, with BE set off at the lower end of the factor, and with the other programs broken up into overlapping ranges. Again, the self-directed activity of pupils specified in the rationales of the EDC and Nimnicht programs is shown, and programmed learning directed by the teacher appears for the BE program.

Factor 3 - Warm Climate - Programs are discriminated statistically significantly on the factor, with three groups being identified as homogeneous. (See Table D-16). Depending on the range examined, the BE program or the BE program plus the Nimnicht program are set off at the upper end of the distribution, and the Comparison, Gotkin, and PE programs are set off at the lower end.

Factor 4 - Pupil Negative Behavior - The two subgroups appear to be created by the Gotkin program being set apart at the high end of the scale.

Table D-15

Multiple Range Test - Florida Climate and Control System - 1970

Factor 1 - Teacher Negative Control   Factor 2 - Pupil Freedom vs Teacher Struc-
                                                        tured Activity

Sponsor	Mean	NSR*	S.D.	N	Sponsor	Mean	NSR*	S.D.	N
Gotkin	55.09		7.53	8	Nimnicht	57.60		3.23	8
Bank St.	54.51		8.07	8	EDC	57.13		3.24	8
Comp.	51.95		6.69	14	Tucson	50.85		1.39	8
PE	50.85		4.22	8	Bank St.	49.85		3.93	8
BE	49.57		4.81	8	Gotkin	48.81		4.16	8
EDC	48.38		4.04	8	PE	47.31		5.32	8
Nimnicht	45.92		5.48	8	Comp.	46.65		4.06	14
Tucson	45.10		4.67	8	BE	44.31		2.03	8

F = 3.10**   F = 14.50**

Table D-16

Multiple Range Test - Florida Climate and Control System - 1970

Factor 3 - Warm Climate   Factor 4 - Pupil Negative Behavior

Sponsor	Mean	NSR*	S.D.	N	Sponsor	Mean	NSR*	S.D.	N
BE	55.51		4.69	8	Gotkin	56.59		5.19	8
Nimnicht	53.95		5.03	8	Nimnicht	51.84		5.30	8
Bank St.	51.44		6.12	8	Bank St.	51.18		7.40	8
EDC	51.28		4.96	8	PE	49.89		3.24	8
Tucson	49.30		6.90	8	EDC	49.79		4.59	8
Comp.	47.58		5.91	14	Comp.	49.07		6.17	14
Gotkin	47.11		4.90	8	Tucson	48.01		4.75	8
PE	45.91		2.61	8	BE	47.69		4.80	8

F = 3.43**   F = 2.27*

(See Table D-16). It seems probable that the position of this program again reflects the one school, cited earlier, in which considerable tension existed.

Factor 5 - Structured Teacher Behavior with Praise - The program which stands highest on this factor is the BE program (Table D-17). In a sense the factor seems to represent two aspects of BE rationale, in which the programmed learning materials structure the behavior of the teacher, and the teacher motivates the pupils by the use of praise. Another of the ranges separates off the Gotkin program at the upper end, following the BE program, presumably because of the extent to which the Gotkin materials structure the behavior of teachers in his classrooms.

Factor 6 - Pupil Conforming Behavior - There are no discriminations between protrams, and the F suggests that variability within the programs is larger than differences between programs (Table D-17).

Factor 7 - Effective Personal Communication - The factor creates three subgroups, and discriminates programs significantly (Table D-18). In one separation the BE program is placed alone at the top; in another the BE and Nimnicht programs are set off together at the top of the range. The Gotkin program is separated at the lower end, perhaps as a consequence of a film that was circulating through classrooms during the days in which observation was being conducted in one of the two cities from which the Gotkin data came. With samples as small as these, such incidental occurrences can have considerable effect on the placement of a group.

Table D-17

Multiple Range Test - Florida Climate and Control System - 1970

Factor 5 - Structured Teacher Behavior   Factor 6 - Pupil Conforming Behavior
with Praise

Sponsor	Mean	NSR*	S.D.	N	Sponsor	Mean	NSR*	S.D.	N
BE	60.60		3.79	8	BE	51.67		5.31	8
Gotkin	51.53		3.45	8	Bank St.	51.38		4.73	8
Tucson	50.05		2.67	8	EDC	51.29		4.73	8
Comp.	49.35		2.84	14	Gotkin	49.98		4.61	8
Nimnicht	49.08		2.78	8	Comp.	49.83		6.17	14
Bank St.	48.33		3.54	8	PE	49.65		5.06	8
EDC	46.85		3.28	8	Tucson	49.33		7.90	8
PE	46.60		5.58	8	Nimnicht	48.02		3.05	8

F = 12.83**   F = .42 N.S.

Table D-18

Multiple Range Test - Florida Climate and Control System - 1970

Factor 7 - Effective, Personal Communication

Sponsor	Mean	NSR*	S.D.	N
BE	58.13		7.11	8
Nimnicht	54.10		4.19	8
PE	49.48		2.50	8
Tucson	49.15		8.32	8
Bank St.	48.93		5.55	8
Comp.	48.60		5.74	14
EDC	47.73		5.30	8
Gotkin	45.83		4.62	8

F = 3.94**

Reciprocal Category System

Factor 1 - Drill with Teacher Acceptance and Pupil Initiation vs Teacher and Pupil Initiation - The BE program is separated from the others both years (Table D-19). This is not surprising since BE teachers spend much of their time interacting with subgroups of pupils working on programmed learning materials, and one of their objectives is to foster a high rate of pupil response. The pattern is not the usual "drill" in that the teachers do talk at greater length than would be true of a rapid-fire drill session, they accept answers, and praise pupils at length.

Factor 2 - Varied Pupil Directed Interaction - The factor discriminates significantly between programs both years, but the order shifts. (See Table D-20). The EDC program stands at or near the upper end of the scale both years, the Nimnicht program is set off at the upper end the second year, and BE is set off at the lower end of the scale the second year. The other programs are not separated from each other. The high standings of the EDC and Nimnicht programs seem reasonable in that in both pupils are freer to interact with each other in various ways. Correspondingly, the programmed learning context in which much of the interaction of teacher and pupils occurs in the BE program places the pupil in the position of responding to the teacher, rather than directing the interaction.

Factor 3 - Warm, Nonevaluative Teaching vs Teacher Criticism - Although the F is not significant either year, the multiple range test separates programs into two groups the second year -- the Tucson program is separated from the rest of the group at the upper end, and the Bank Street and Comparison groups are separated off at the lower end (Table D-21). The high standing of the Tucson program appears to relate to its commitment to a nonevaluative climate.

Table D-19

Multiple Range Test - Reciprocal Category System

Factor 1 - Drill with Teacher Acceptance and Pupil Initiation vs Teacher and Pupil Initiation

Sponsor	1969 Mean	NSR*	S.D.	N	Sponsor	1970 Mean	NSR*	S.D.	N
BE	56.31		5.25	8	BE	59.97		3.41	8
PE	52.37		2.66	8	Gotkin	51.45		3.73	8
Gotkin	49.97		3.74	8	PE	51.11		5.96	8
Bank St.	49.76		3.38	8	Comp.	49.94		6.43	13
Comp.	48.95		4.07	14	Bank St.	48.45		4.49	6
Tucson	48.49		3.17	8	EDC	46.17		6.45	6
EDC	48.24		1.34	8	Tucson	46.10		6.85	8
Nimnicht	46.13		5.60	8	Nimnicht	44.84		7.00	8

$F = 5.12 \quad p < .01$ $\qquad\qquad$ $F = 5.39 \quad p < .01$

Table D-20

Multiple Range Test - Reciprocal Category System

Factor 2 - Varied Pupil Directed Interaction

Sponsor	1969 Mean	NSR*	S.D.	N	Sponsor	1970 Mean	NSR*	S.D.	N
EDC	55.26		3.58	8	Nimnicht	57.16		6.52	8
Gotkin	52.62		5.86	8	EDC	53.23		3.85	6
BE	50.39		3.69	8	Tucson	52.48		6.00	8
Nimnicht	49.67		4.43	8	Gotkin	51.60		3.53	8
PE	48.89		3.18	8	Comp.	48.31		7.00	13
Comp.	48.76		4.66	14	PE	47.75		5.68	8
Bank St.	48.12		4.27	8	Bank St.	47.10		4.17	6
Tucson	47.43		6.50	8	BE	43.87		6.92	8

$F = 2.58 \quad p < .05$ $\qquad\qquad$ $F = 4.06 \quad p < .01$

Table D-21

Multiple Range Test - Reciprocal Category System

Factor 3 - Warm, Nonevaluative Teaching vs Teacher Correction

Sponsor	Mean	1969 NSR*	S.D.	N	Sponsor	Mean	1970 NSR*	S.D.	N
Tucson	55.69		7.80	8	Tucson	55.40		6.41	8
BE	54.47		5.26	8	BE	52.98		3.54	8
Bank St.	49.28		6.70	8	EDC	51.25		6.25	6
EDC	49.16		9.81	8	Nimnicht	50.50		6.82	8
Comp.	48.77		10.12	14	PE	48.73		7.62	8
PE	48.09		5.65	8	Gotkin	48.73		6.45	8
Nimnicht	46.25		11.48	8	Comp.	47.26		5.89	13
Gotkin	46.13		5.25	8	Bank St.	45.69		3.91	6

F = 1.45 N.S.   F = 2.15 N.S.

Table D-22

Multiple Range Test - Reciprocal Category System

Factor 4 - Teacher Amplification in Extended Talk

Sponsor	Mean	1969 NSR*	S.D.	N	Sponsor	Mean	1970 NSR*	S.D.	N
Gotkin	51.80		3.34	8	PE	52.15		6.30	8
PE	50.85		1.69	8	Comp.	51.70		5.23	13
Tucson	50.42		3.46	8	Tucson	51.60		6.10	8
Comp.	49.94		3.03	14	BE	50.44		3.78	8
Bank St.	49.32		2.53	8	Gotkin	49.83		5.33	8
BE	49.22		5.82	8	Bank St.	49.28		5.86	6
Nimnicht	49.14		3.87	8	EDC	48.08		6.95	6
EDC	48.62		3.16	8	Nimnicht	46.31		4.94	8

F = 0.73 N.S.   F = 1.07 N.S.

The low standing of the Bank Street program the second year probably relates to the situation of one of its schools, as cited in the body of the report.

Factor 4 - Teacher Amplification in Extended Talk - The factor does not discriminate programs, and the F suggests that variability within programs and variability between programs are essentially equal (See Table D-22).

Factor 5 - Pupil Talk - The factor does not discriminate between programs either year, with the F suggesting that variability within programs is as large or larger than between programs. (See Table D-23).

Factor 6 - Teacher Initiation with Pupil Interruption vs Teacher Question - Programs are discriminated significantly both years, with two groups the first year and three the second (Table D-24). The BE program is separated off at the lower end both years; the Tucson program was set off at the upper end the second year. The rationale of the BE program limits extended teacher initiation in favor of soliciting pupil responses.

Factor 7 - Teacher Direction vs Pupil Initiation - The factor did not discriminate significantly the first year but discriminates highly significantly between programs the second year (Table D-25). Although four ranges are identified, the gaps between means suggest three groupings: BE, Comparison and Gotkin at the upper end; PE, Bank Street and Tucson in the midrange; and EDC and Nimnicht at the lower end. The extreme positions appear to agree with program rationale. Teacher direction of learning is a central aspect of the programming of learning (BE), and minimal direction of pupils seems equally central to the Nimnicht classrooms in which pupils are expected to learn from self-teaching equipment and EDC classrooms in which the child is exposed to rich materials and encouraged to follow his interests to a large degree.

Table D-23

Multiple Range Test - Reciprocal Category System

Factor 5 - Pupil Talk

Sponsor	1969				Sponsor	1970			
	Mean	NSR*	S.D.	N		Mean	NSR*	S.D.	N
EDC	52.04		6.01	8	Nimnicht	53.43		5.35	8
BE	51.78		9.27	8	BE	52.89		7.51	8
Tucson	51.75		8.40	8	Tucson	52.37		7.08	8
Bank St.	49.84		7.34	8	PE	50.45		7.69	8
Comp.	49.32		7.09	14	EDC	49.26		7.80	6
PE	49.25		6.20	8	Bank St.	48.64		10.41	6
Gotkin	49.18		12.62	8	Comp.	46.96		8.51	13
Nimnicht	47.50		5.80	8	Gotkin	46.13		3.70	8

F = 0.33 N.S.   F = 1.18 N.S.

Table D-24

Multiple Range Test - Reciprocal Category System

Factor 6 - Teacher Initiation with Pupil Interruption vs Teacher Question

Sponsor	1969				Sponsor	1970			
	Mean	NSR*	S.D.	N		Mean	NSR*	S.D.	N
Nimnicht	54.55		7.0T	8	Tucson	55.90		5.27	8
Comp.	51.69		4.62	14	Nimnicht	53.24		4.73	8
Tucson	50.93		11.75	8	Comp.	52.51		9.43	13
PE	49.82		6.33	8	EDC	51.52		4.72	6
Bank St.	49.71		5.57	8	Bank St.	49.43		4.92	6
EDC	49.40		7.47	8	PE	48.25		4.67	8
Gotkin	49.20		6.30	8	Gotkin	47.25		5.30	8
BE	39.73		4.57	8	BE	39.85		5.11	8

F = 3.22 $p < .01$ F = 5.10 $p < .01$

Table D-25

Multiple Range Test - Reciprocal Category System

Factor 7 - Teacher Direction vs Pupil Initiation

Sponsor	1969				Sponsor	1970			
	Mean	NSR*	S.D.	N		Mean	NSR*	S.D.	N
Gotkin	51.96		3.70	8	BE	55.48		4.03	8
PE	51.39		4.13	8	Comp.	53.80		5.97	13
Comp.	51.31		5.68	14	Gotkin	53.33		4.87	8
Bank St.	49.97		4.42	8	PE	49.25		8.13	8
BE	49.92		3.76	8	Bank St.	49.09		6.38	6
Tucson	49.34		5.75	8	Tucson	48.27		5.10	8
Nimnicht	47.65		4.89	8	EDC	43.08		5.78	6
EDC	46.62		4.75	8	Nimnicht	42.27		7.86	8

F = 1.30 N.S.   F = 4.97   p < .01

Table D-26

Multiple Range Test - Reciprocal Category System

Factor 8 - Supportive Pupil Talk

Sponsor	1969				Sponsor	1970			
	Mean	NSR*	S.D.	N		Mean	NSR*	S.D.	N
EDC	54.21		7.85	8	Nimnicht	55.45		8.28	8
Gotkin	53.38		8.82	8	Gotkin	54.59		7.60	8
PE	51.67		7.63	8	Tucson	51.04		4.61	8
Bank St.	50.92		7.29	8	EDC	50.89		3.70	6
BE	49.92		5.54	8	PE	50.25		7.68	8
Tucson	48.92		5.42	8	BE	48.92		6.57	8
Comp.	47.62		1.58	14	Comp.	48.36		4.02	13
Nimnicht	47.50		1.41	8	Bank St.	46.50		1.22	6

F = 1.58 N.S.   F = 2.05 N.S.

Factor 8 - Supportive Pupil Talk - As mentioned in the body of the report, this factor appears to be a reflection of the achievement of program objectives directed at fostering supportive social behavior among pupils. The discrimination between programs did not result in a significant F either year (Table D-26). The multiple range test separated off the EDC program at the upper end of the scale the first year, and the Gotkin and Nimnicht programs the second year. Bank Street and Comparison classrooms were separated off at the lower end of the scale the second year.

This measure, made up of pupil talk, may be subject to the bias cited in the Procedure section that some programs more than others organize pupil activity in such a way that their talk is audible. Pupils are likely to be more audible in BE classrooms, for instance, and less so in EDC classrooms.

Cognitive Taxonomy

Factor 1 - Highly Focused Learning Tasks + The multiple range tests for each of the two years place the Becker-Engelmann program at the top of the scale (Table D-27). There is no consistency in the programs separated at the lower end of the factors the two years. The position of the BE program seems in line with its orientation toward fostering the development of skills. Other programs shift extensively in their relative rank order in the remainder of the range so that probably only the BE position can be taken as meaningful.

Factor 2 - Narrow vs Broad Answer - Although the positive pole is labeled "narrow" it represents a high cognitive level; the negative pole appears to be less subject-matter oriented, and at a lower cognitive level, but no one answer would be expected. Each year, programs were separated into two groups, but the F's were not significant either year. (See Table D-28). The first year, the BE program was separated off at the upper end of the scale, the second year it was a part of the top group along with Comparison classrooms and the Nimnicht program. Both years, the EDC program was separated off at the lower end of the scale. These separations seemed to fit satisfactorily with the concern for fostering subject matter skills, which is typical of BE classrooms, and which has been recognized as a primary concern of Comparison classrooms. Nimnicht classrooms are also high on this factor and appear to support the development of subject matter skills, although they do so in a different way than the other two programs. The EDC program, in contrast, does not make that sort of skill development a central issue in the early years.

Factor 3 - Moderately Focused Learning Tasks - The factor discriminates significantly between programs the second year, and approaches significance the first year (Table D-29). Two groups are set off each year, with BE at the

Table D-27

Multiple Range Test - Cognitive Taxonomy

Factor 1 - Highly Focused Learning Tasks

Sponsor	1969 Mean	NSR*	S.D.	N	Sponsor	1970 Mean	NSR*	S.D.	N
BE	54.78		5.30	8	BE	57.06		5.37	8
PE	53.23		6.85	8	Gotkin	52.14		4.13	8
Bank St.	50.96		7.60	8	Nimnicht	51.80		5.26	8
EDC	49.94		4.93	8	Comp.	48.99		6.10	13
Tucson	48.73		7.17	8	Bank St.	48.95		4.92	6
Gotkin	48.51		3.77	8	PE	48.38		3.99	8
Nimnicht	48.19		5.14	8	EDC	46.77		2.82	6
Comp.	47.33		7.02	14	Tucson	46.30		4.55	8

F = 1.60   F = 3.98**

Table D-28

Multiple Range Test - Cognitive Taxonomy

Factor 2 - Broad vs Narrow Answer

Sponsor	1969 Mean	NSR*	S.D.	N	Sponsor	1970 Mean	NSR*	S.D.	N
BE	53.58		4.79	8	Nimnicht	51.99		3.95	8
Bank St.	52.78		5.79	8	Comp.	51.31		3.26	13
PE	51.00		5.76	8	BE	50.41		3.27	8
Comp.	50.66		5.52	14	Bank St.	50.40		3.04	6
Tucson	50.18		4.81	8	Tucson	49.93		2.45	8
Gotkin	48.29		3.34	8	Gotkin	49.54		2.55	8
Nimnicht	47.64		2.89	8	PE	48.28		4.24	8
EDC	47.14		7.40	8	EDC	46.32		2.43	6

F = 1.61   F = 2.21*

Table D-29

Multiple Range Test - Cognitive Taxonomy

Factor 3 - Moderately Focused Learning Tasks

Sponsor	1969 Mean	NSR*	S.D.	N	Sponsor	1970 Mean	NSR*	S.D.	N
BE	53.91		5.98	8	BE	54.20		3.04	8
PE	53.28		5.69	8	Gotkin	52.31		3.78	8
EDC	51.58		6.79	8	Nimnicht	51.23		2.34	8
Nimnicht	50.39		5.30	8	Bank St.	50.70		5.20	6
Bank St.	50.38		5.74	8	PE	49.94		4.61	8
Gotkin	48.05		6.26	8	Tucson	47.98		3.70	8
Tucson	47.33		4.44	8	Comp.	47.72		4.67	13
Comp.	46.04		6.66	14	EDC	47.53		2.08	6

$F = 2.13$ $F = 3.15**$

Table D-30

Multiple Range Test - Cognitive Taxonomy

Factor 4 - Phonics

Sponsor	1969 Mean	NSR*	S.D.	N	Sponsor	1970 Mean	NSR*	S.D.	N
BE	55.60		4.07	8	BE	55.13		6.78	8
PE	52.13		3.44	8	Bank St.	53.07		5.53	6
Tucson	49.96		4.16	8	Tucson	50.11		4.81	8
Nimnicht	49.36		5.44	8	Comp.	49.79		4.92	13
EDC	48.99		6.00	8	Gotkin	49.10		5.30	8
Comp	48.89		4.36	14	PE	48.76		4.92	8
Bank St.	48.55		4.84	8	EDC	47.90		2.24	6
Gotkin	47.26		3.76	8	Nimnicht	47.01		3.69	8

$F = 2.70*$ $F = 2.21*$

upper end both years, accompanied by PE the first year and Gotkin the second year. Comparison classrooms are set off at the lower end both years, joined by EDC and Tucson the second year. The placement of the BE program appears to reflect its emphasis on teaching skills, and related broader generalizations such as classification. The placement of EDC classrooms also seems reasonable in relation to their lower emphasis on teaching skills. The position of Comparison classrooms seems surprising in relation to the expectation that teaching of skills would be emphasized there.

It is probably relevant to mention again the caution cited in the Procedure section, that some programs are likely to be more completely represented on tape than others -- probably BE would be better represented and EDC less well represented in terms of the proportion of pupil activity which is oral, and the proportion of what is oral that is audible. To some degree, the placement of programs on the data of the Cognitive Taxonomy is probably biased in this respect.

Factor 4 - Phonics - Programs were discriminated significantly, with two groups created both years. The BE program was set off at the upper end of the dimension both years (Table D-30), and again, the BE emphasis on skill teaching seems relevant.

Factor 5 - Information Giving and Receiving - Programs were discriminated significantly the first year but not the second (Table D-31). In the first year's data, the PE program is set off at the upper end of the factor, the Tucson at the lower end, joined by Bank Street and Gotkin set off by other ranges. Although Parent Education is not primarily a classroom oriented program, it has had a varying emphasis on a Piagetian theory of child development. Pupils' establishment of an adequate base of information (as well as higher level cognitive activities) is relevant to that theoretical orientation. Perhaps the

Table D-31

Multiple Range Test - Cognitive Taxonomy

Factor 5 - Information Giving and Receiving

Sponsor	1969 Mean	NSR*	S.D.	N	Sponsor	1970 Mean	NSR*	S.D.	N
PE	54.38		5.43	8	BE	51.81		6.65	8
Nimnicht	51.70		5.80	8	Comp.	51.33		4.16	13
BE	51.63		5.65	8	Gotkin	50.18		4.50	8
EDC	51.60		6.95	8	EDC	49.80		6.63	6
Comp.	49.29		6.30	14	Bank St.	49.75		4.21	6
Gotkin	47.41		5.12	8	Nimnicht	48.80		5.18	8
Bank St.	47.38		6.73	8	Tucson	48.40		5.53	8
Tucson	43.69		4.78	8	PE	48.05		5.96	8

F = 2.61* F = .56 N.S.

Table D-32

Multiple Range Test - Cognitive Taxonomy

Factor 6 - Complex Thinking

Sponsor	1969 Mean	NSR*	S.D.	N	Sponsor	1970 Mean	NSR*	S.D.	N
BE	53.25		5.28	8	Nimnicht	52.56		5.02	8
Bank St.	51.25		3.94	8	Gotkin	51.48		6.35	8
Nimnicht	50.14		5.40	8	Comp.	50.80		4.87	13
PE	49.99		5.28	8	PE	50.65		6.21	8
Comp.	49.71		4.75	14	BE	49.40		4.14	8
EDC	49.66		3.73	8	EDC	49.15		3.71	6
Gotkin	49.34		5.76	8	Bank St.	48.80		2.16	6
Tucson	48.86		5.62	8	Tucson	48.69		5.63	8

F = .62 N.S. F = .59 N.S.

failure of programs to separate the second year (and the narrowed range) reflects increasing concern of all sponsors with this objective, although this interpretation is speculative.

Factor 6 - Complex Thinking - This factor did not discriminate between programs either year, and the F each year suggests that variability within programs is as great or greater than differences between programs (Table D-32). This seems a surprising finding in the light of differences between programs in their emphasis on inquiry learning. It seems possible that the program bias on audibility cited earlier may have the effect of artificially minimizing program differences on this factor.

Appendix E

Table E-1

Reciprocal Category Measures

Variable	Description
1	Teacher warms, informalizes the climate. The sum of column 1.
2	Teacher accepts. The sum of column 2.
3	Teacher amplifies. The sum of column 3.
4	Teacher elicits. The sum of column 4.
5	Teacher responds. The sum of column 5.
6	Teacher initiates. The sum of column 6.
7	Teacher direct. The sum of column 7.
8	Teacher corrects. The sum of column 8.
9	Teacher cools, formalizes. The sum of column 9.
10	Pupil elicits. The sum of 14.
11	Pupil responds. The sum of 15.
12	Pupil initiates. The sum of 16.
13	Pupil directs. The sum of 17.
14	Pupil corrects. The sum of 18.
15	Confusion (does not include uncodeable tape). The sum of column 20.
16	Teacher talk, percent. The sum of columns 1-9 divided by columns 1-9 plus columns 11-19.
17	Teacher acceptance-rejection, percent. The sum of columns 1, 2, and 3 divided by columns 1, 2, and 3 plus 8 and 9.
18	Pupil initiation. The sum of column 16 divided by total student talk, the sum of columns 11-19.

Table E-1 - Continued

Variable	Description
19	Student response to teacher. Rows 1-9 for column 15, divided by total student talk.
20	Pupil-pupil talk. The sum of rows 11-19 for columns 11-19.
21	Teacher extended indirect. The sums of the cells in rows 1, 2, and 3 for columns 1, 2, and 3.
22	Teacher extended direct. The sum of the cells in rows 7-9 for columns 7-9.
23	Teacher revised I/D. This measure involves teacher indirect response (rows 11-19 for columns 1-3), and teacher direct response (rows 11-19 for columns 7-9). The percentage is made up of indirect response divided by indirect response plus direct response.
24	Pupil positive participation, percent. Positive participation divided by positive participation plus negative participation (rows 15, 16, for columns 11-13/rows 15, 16 for columns 11-13, plus rows 15, 16 for columns 17-19).
25	Pupil revised I/D_1. This measure includes pupil extended indirect (rows 11-19 for columns 11-13); and pupil extended direct (rows 11-19 for columns 17-19); with pupil extended indirect divided by pupil extended direct, plus pupil extended indirect.
26	Teacher narrow question. The sum of the 4-15 cell.
27	Teacher broad question. The sum of the 4-16 cell.
28	Pupil broad question. The sum of the 14-16 cell.
29	Pupil substantive interruption. The row 6, columns 14-16 cells.
30	Pupil direct interruption. The row 6, columns 17-19 cells.
31	Total pupil interruption. The total of pupil substantive interruption plus pupil direct interruption.
32	Pupil question, teacher question. The 14-4 cell.
33	Pupil question, teacher response. The 14-5 cell.