DOCUMENT RESUME

ED 034 724 SP 003 399

AUTHOR Hill, Russell A.; Medley, Donald M.
TITLE Change in Behaviors of First Year Intern Teachers.
PUB DATE Feb 68
NOTE 17p.; Paper presented at the American Educational Research Association meeting, Chicago, February, 1968

EDRS PRICE EDRS Price MF-$0.25 HC-$0.95
DESCRIPTORS *Beginning Teachers, *Behavior Change, Student Teacher Relationship, *Teacher Behavior, Teacher Improvement, *Teacher Interns

ABSTRACT

A study measured changes in the teaching behavior of 70 first-year Temple University interns teaching in metropolitan Philadelphia secondary schools. Subjects, selected from 110 on the basis of subject and grade level, had received 6 weeks of graduate training in the summer prior to placement; they participated throughout the year in seminar classes while receiving close University supervision. Each of two observational instruments--Flanders' Interaction Analysis (IA) and Medley's Observation Schedule and Record 4-verbal (OScAR 4V)--were used by one of two observers who visited the same teacher at the same time; interns where observed four times, twice in early February and twice in late May. The differences between the pairs of observation were expected to reflect positive change over the hypothetical period of accelerating growth. Significant differences were found on 14 IA scales and on 15 OScAR scores; OScAR scores were also rescaled to form orthogonal contrasts, and 14 of these new scales indicated significant changes in teaching behavior. Overall results: In May the intern teachers were describing more, using more divergent questions and less convergent questions, and being less evaluative and more neutral in their responses; there appeared to be a shift from direct student response to student initiated responses with student responses more often accepted or neutrally evaluated rather than praised or criticized. (SP 003 398 is a related document.) (JS)

during each visit. Observers sampled behavior for a period of twenty-eight (28) minutes during the period.

The observers were twelve supervisors from the Intern Teaching Program who were divided into six teams of two. (Table I) One member of each team was trained in IA observation, the other in OScAR 4V observation. IA observers received approximately twelve (12) hours of training prior to the January observation and three (3) hours of refresher training before the May observations. OScAR observers received eight (8) hours of initial training and two (2) hours of refresher training.

SCORING

The data collected was scored for individual teachers. Three types of scores were obtained. The ten IA categories were scored upon forty-four (44) different scales. These scales included many of the traditional scores reported in the literature (Amidon et al, 1963) as well as other scores selected for their apparent significance and interpretability.

Forty-two (42) independent scores were obtained from the OScAR 4V categories. In addition, a new procedure for scoring OScAR 4V was introduced. The procedures followed involved selecting linear contrasts among categories. (See Winer, 1962, page 70) OScAR categories were given positive, negative or 0 weights. Weights were selected to meet the criteria of orthogonality; i.e., weights multiplied by unity sum to zero for every score, and the cross-products of the weights of any two scales sum to zero. Forty-two (42) scales were chosen from the set of

A.E.R.A. PRESENTATION - Chicago, February 1968

CHANGE IN BEHAVIORS OF FIRST YEAR INTERN TEACHERS

Russell A. Hill, Temple University
Donald M. Medley, Educational Testing Services

This study measures changes in the teaching behavior of seventy first year intern teachers. Two observational instruments were used: Flanders' Interaction Analysis technique (IA) and Medley's Observation Schedule and Record 4-verbal (OScAR 4V). Each of these instruments was used by one of two observers who visited the same teacher at the same time. The seventy interns were observed four times, twice in early February and twice in late May. The difference between the pairs of observation was considered to be the change. Significant differences were found on fourteen (14) IA scales and on fifteen (15) OScAR scores. In addition, the OScAR scores were rescaled to form orthogonal contrasts and fourteen (14) of these new scales indicated significant changes in teaching behavior.

SAMPLE DESCRIPTION

Intern teachers observed were members of the Intern Teaching Program at Temple University. Members of the Program are selected liberal arts graduates who have enrolled in a graduate teacher training program. The interns receive six weeks of training in their initial summer and are placed in full-salaried, full-time teaching positions in the fall. During their first year they participate in seminar classes and receive close University supervision. The interns continue to receive supervision and to participate in course work for two years following their initial

placement. Ultimately, they earn a Master's degree in Education and full State certification.

Supervision experience with interns has led to the classification of stages of intern development. The first three months of teaching appear to be a highly emotional period of scrambling in the classroom. The next three months is a period of quiet adjustment and development of basic classroom routines. The final three months of the year are perceived as a period of accelerating growth in educational planning and instructional skill. This study focused upon these last three months. Initial observations were made in early February (the weeks of January 24 and February 4). Final observations were made in late May (the weeks of May 16 and May 27). Thus, measurements span this hypothetical period of accelerating growth and should reflect a positive change in teaching behavior.

All interns observed were in metropolitan Philadelphia and ranged in location from the most favored suburbs to hard core ghetto areas. The individual interns were selected from a larger pool of one hundred ten (110) first year interns on the basis of subject and grade level. They were divided into six groups by subject and by junior or senior high schools. (Table I) Of the seventy (70) first year interns observed, thirty-six (36) were at the senior high school level and thirty-four (34) were at the junior high school level. Four (4) subject areas were observed: Mathematics, Science, English and Social Studies. Forty-three (43) were men and twenty-seven (27) were female.

TABLE I

SCHEDULE DESIGN FOR OBSERVATION OF FIRST YEAR INTERNS AT TEMPLE UNIVERSITY

Observer Team	Level	Subject #1	Subject #2
1.	Jr. H.S.	English (6)*	Mathematics (6)
2.	Jr. H.S.	English (6)	Science (5)
3.	Jr. H.S.	Soc.Studies(6)	Science (5)
4.	Sr. H.S.	English (6)	Mathematics (6)
5.	Sr. H. S.	English (6)	Science (6)
6.	Sr. H. S.	Soc.Studies(6)	Science (6)

*Number of interns observed in subject at level

PROCEDURES FOR DATA COLLECTION

... Each intern was observed four times by two observers. Interns were notified and observation visits were scheduled ahead of time. Arrangements were made to insure that interns were observed teaching the same class of students during each of the four visits.

Two different observational instruments were used; Flanders' Interaction Analysis technique (IA) (Amidon et al, 1963) which divides student and teacher behavior into ten categories and the Observation Schedule and Record 4 verbal (OScAR 4V) (Medley, 1962) which involves tallying teacher and pupil initiated statements and interchanges into forty-two (42) mutually exclusive cells. Teaching behavior was recorded simultaneously on each instrument to produce two independent records of teacher behavior during

during each visit. Observers sampled behavior for a period of twenty-eight (28) minutes during the period.

The observers were twelve supervisors from the Intern Teaching Program who were divided into six teams of two. (Table I) One member of each team was trained in IA observation, the other in OScAR 4V observation. IA observers received approximately twelve (12) hours of training prior to the January observation and three (3) hours of refresher training before the May observations. OScAR observers received eight (8) hours of initial training and two (2) hours of refresher training.

SCORING

The data collected was scored for individual teachers. Three types of scores were obtained. The ten IA categories were scored upon forty-four (44) different scales. These scales included many of the traditional scores reported in the literature (Amidon et al, 1963) as well as other scores selected for their apparent significance and interpretability.

Forty-two (42) independent scores were obtained from the OScAR 4V categories. In addition, a new procedure for scoring OScAR 4V was introduced. The procedures followed involved selecting linear contrasts among categories. (See Winer, 1962, page 70) OScAR categories were given positive, negative or 0 weights. Weights were selected to meet the criteria of orthogonality; i.e., weights multiplied by unity sum to zero for every score, and the cross-products of the weights of any two scales sum to zero. Forty-two (42) scales were chosen from the set of

possible contrasts on the basis of interpretability.

Thus, the scoring procedures yielded forty-four (44) IA scores, forty-two (42) OScAR raw scores, and forty-two (42) OScAR orthogonal scales for a total of one hundred twenty-eight (128) individual scores for each of the seventy(70) intern teachers.

DATA ANALYSIS

Individual scores were subjected to analysis of variance. The overall scores were contrasted with subject scores. Reliabilities were estimated for the two observational techniques. The sensitivity of each of the two techniques to teacher behavior change from February to May was determined. (Table II)

TABLE II

DESIGN FOR ANALYSIS OF VARIANCE FOR OBSERVATIONS OF TEMPLE INTERNS BY IA AND OScAR

SOURCE OF VARIATION	D.F.
Change	1
Visit	1
Change x visit	1
Observer	5
Change x observer	5
Visit x observer	5
Change x visit x observer	5
Subject (in observer)	6
Change x subject (in observer)	6
Visit x subject (in observer)	6
Change x visit x subject (in observer)	6
Teacher (in subject and observer)	58
Change x teacher (in subject & observer)	58
Visit x teacher (in subject & observer)	58
Residual	58
Total Variation	279

RESULTS

Only the scores and scales indicating significant change are reported in this paper. Initially, each set of scores is reported and interpreted separately. An interpretation utilizing the significant changes of all the scores and scales is reserved until the conclusion.

IA.

Fourteen (14) Interaction Analysis scores indicated significant changes in the teaching behavior of the intern. The scores seem to be easily interpretable and non-contradicting. (Table III) The following statements can be made regarding significant changes:

1. Student initiated responses (9's) increased dramatically while direct student responses (8's) decreased. This shift seemed to have a major influence across all scores. That is, most of the scores containing a heavy weighting of 9's increased (all 9 of the positive changes were so weighted) and several of the scores that decreased were weighted with 8's (4 out of 6 negative changes were so weighted).
2. There was a decrease in the use of praise by the teacher. (2's)
3. There was a shift from convergent (48's) to divergent (49's) questions by the teacher.
4. There was a decrease in the area known as the content-cross.
5. Increased student talk was followed by both an increase in the amount of teacher acceptance (3's) and in no

TABLE III

SIGNIFICANT CHANGE BY INTERNS ON INTERACTION ANALYSIS SCORES BY DIRECTION

INCREASING FREQUENCY (+)

	Name of Score	Mean Freq. Pre	Mean Freq. Post	Mean Change	Items tabulated for scale (Matrix Cells)
1.	Student initiates	35.79	50.54	14.76**	91 thru 910. 19,29,39, 49,59,69,79,89,109
2.	Student initiations accepted	2.00	3.74	1.74**	91, 93
3.	Student initiations not evaluated	5.09	6.51	1.42**	94,95,96
4.	Divergent questions	2.87	3.88	1.01**	49
5.	Direct stimulation	0.43	0.69	0.26*	69
6.	Student initiation following teacher talk	2.16	2.73	0.57*	19, 29, 39
7.	Student initiation following teacher lecture	3.39	4.66	1.27**	59
8.	Indirect vs. direct response to student initiation	3.02	4.99	1.97**	+91,+92,+93,-96,-97

DECREASING FREQUENCY (-)

9.	Praise	22.27	20.22	-2.75*	21 thru 210.12,32,42, 52,62,72,82,92,102
10.	Direct student response	68.48	57.91	-10.56*	81 thru 810.18,28,38, 48,58,68,78,98,108
11.	Content Cross	75.33	69.98	- 5.35*	14,15,24,25,34,35,41, 42,43,46 thru 53.56 thru 510.64,65,74,75, 84,85,94,104,105
12.	Student response, praised	6.18	5.04	-1.14*	82
13.	Direct student response not evaluated	7.85	6.13	-1.72**	84,85,86
14.	Convergent questions	20.01	16.94	-3.07**	48

(** (.01 level of sig.
(* (.05 level of sig.

evaluation by the teachers.

6. Student talk seemed to increase following all teacher talk, and increased most specifically following lectures, divergent questions and directions.

OScAR RAW SCORES

Significant changes were reflected on fourteen (14) OScAR raw scores. These significant scores are presented in Table IV. The following statements seem to encapsulate their significance.

1. Teacher describing increased both in frequency and in length.
2. Teacher questions became shorter.
3. Elaborating interchanges decreased.
4. Divergent interchanges increased.
5. There was a consistent shift within all interchanges from both positive and negative termination of interchange toward "acceptance" as defined in OScAR. (Acceptance is used in the sense of perfunctory acknowledgment in OScAR.)
6. Teachers lectured for a longer period of time once they began. (Inspection of the data suggests that the variance of one observer team has unduly influenced this score. Therefore, it should be considered questionable.

TABLE IV

SIGNIFICANT CHANGE BY INTERNS ON OSCAR RAW SCORES BY DIRECTION

NAME OF SCORE	Mean Freq. Pre	Mean Freq. Post	Mean Change
INCREASE IN FREQUENCY (+)			
1. Describing statement; initiating	3.45	4.41	.96*
2. Describing statement; continuing	1.48	4.73	3.25**
3. Informing; continuing	23.12	28.58	5.46*
4. Divergent interchanges; accepted	.99	2.14	1.16**
5. Convergent interchanges; accepted	7.25	9.81	2.56*
6. Pupil interchanges; accepted	3.88	5.06	1.19*
DECREASE IN FREQUENCY (-)			
1. Problem structuring; continuing	12.14	6.39	-5.74**
2. Elaborating; approved	4.72	3.11	-1.61**
3. Elaborating; neutrally rejected	.96	.57	- .39*
4. Elaborating; criticized	.46	.17	- .29**
5. Convergent interchanges; approved	11.13	8.20	-2.93**
6. Convergent interchanges; criticized	.93	.44	- .49*

(** (.01 level of sig.)

(* (.05 level of sig.)

OScAR ORTHOGONAL SCALES

Significant changes in intern behavior were indicated by fourteen (14) of the forty-two (42) orthogonal scales. The scales indicating significant changes are presented in Table V. The table indicates weights by category used to construct each scale. The amount and direction of change is also indicated.

Scales based upon contrasting types of pupil and teacher statements (Table V-A - scales 1-5) supported the conclusion suggested by the raw scores and offer little new information.

Scale #6; Cohesion indicates a proportional shift from Elaborating Interchanges to Divergent and Convergent Interchange. Scales 7 thru 9 support the overall shift from highly negative and positive evaluation toward neutral acceptance or neutral rejection of interchanges.

Scales 10 thru 14 do offer new insights. The following statements may be made:

1. Positive and negative evaluation of elaborating interchange increased much more than 4 other interchanges; Scale #10 - Cohesion x Feedback.
2. Both acceptance of divergent interchanges and rejection of convergent interchanges increased; Scale #11 - Divergent x Valence.
3. There was more non-substantive pupil initiation; Scale #12 - Pupil initiation

TABLE V

ORTHOGONAL OScAR SCORES INDICATING SIGNIFICANT CHANGE IN INTERN BEHAVIOR

A. SIGNIFICANT SCALES BASED ON CONTRASTING TYPES OF PUPIL AND TEACHER STATEMENTS

		STATEMENTS								
		Direction		Describing		Informing		Problem Structuring		
	Name of Scale	I	C	I	C	I	C	I	C	Change
1.	Suspense					-1	-1	+1	+1	-11.52**
2.	Problem Complexity							-1	+1	- 6.82**
3.	Lecturing					-1	+1			+ 4.43*
4.	Continued Describing			-1	+1					+ 2.29*
5.	Management	+1	+1	+1	+1					+23.86*

B. SIGNIFICANT SCALE BASED ON CONTRASTING TYPES OF ENTRY

		ENTRY			
	Names of Scale	Divergent	Elaborating	Convergent	Change
6.	Cohesion	-½	+1	-½	-3.58*

C. SIGNIFICANT SCALES BASED ON CONTRASTING TYPES OF EXITS

		EXIT						
	Name of Scale	Support	Approve	Accept.	Not Eval.	Neut. Reject.	Crit.	Change
7.	Feedback	+½	+½	-1	-1	+½	+½	-8.50**
8.	Valence	+1	+1	0	0	-1	-1	-4.82**
9.	Positivity	+1	-1	+1	-1	+1	-1	+8.97**

(** (.01 level of sig.)
(* (.05 level of sig.)

TABLE V (cont.)

D. SIGNIFICANT SCALES BASED ON INTERACTION OF ENTRIES AND EXITS FROM INTERCHANGES

NAME OF SCALE	INTERCHANGES: DIVERGENT SP	AP	AC	N-EV	REJ	CRT	ELABORATING SP	AP	AC	N-EV	REJ	CRT	CONVERGENT SP	AP	AC	N-EV	REJ	CRT	CHANGE
10. Cohesion x Feedback	-½	-½	+½	+½	-½	-½	+½	+½	-1	-1	+½	+½	-½	-½	-½	+½	-½	-½	+1.83**
11. Divergent x Valence	+1	+1			-1	-1							-1	-1			+1	+1	+3.39**
12. Divergent x Positivity	+1	-1	+1	-1	+1	-1							-1	+1	-1	+1	-1	+1	-4.94**

E. SIGNIFICANT SCALES CONTRASTING TYPES OF NON-SUBSTANTIVE INTERCHANGES

	NON-SUBSTANTIVE INTERCHANGES: Teacher + Initiated	Teacher - Initiated	Pupil + Initiated	Pupil - Initiated	Change
13. Pupil initiation	-1	-1	+1	+1	+1.91**
14. Encourage pupil initiation	+1	-1	-1	+1	-1.84**

(** (.01 level of sig.)

(* (.05 level of sig.)

SUMMARY OF RESULTS

The three sets of scores tend to support each other and indicate a clear directional change in the behavior of the interns. In some cases, one set of scores tends to give further illumination to a second set of scores. For example, the three kinds of teacher interchanges and the six possible responses greatly amplify the single I.A. question category. In a few cases, there are apparent contradictions between scores. Table VI summarizes the inter-relations of the scores.

It should be remembered that the OScAR Raw Scores and the OScAR Orthogonal Scales are dealing with the same data and thus are subject to the same error. In a sense they are a reamplification of the measures.

DISCUSSION

Difference between scores may be attributed to two factors: (1) difference in the structure of the instruments, and (2) observer reliability.

The instruments differ in several ways. The difference in categories is probably the source of the greatest difference; i.e., IA makes a distinction between direct student responses to questions. OScAR does not. OScAR takes the natural language of formations of statements and full interchanges as its units of measure. IA arbitrarily sets the three second interval and every change in category as its unit of measure. These differences operate together to insure richer and more varied data about the same classroom.

TABLE VI

SUMMARY OF THREE SETS OF SCORES INDICATING SIGNIFICANT CHANGE IN INTERN BEHAVIOR

	TYPE OF SCALE		
	I.A.	OScAR RAW	OScAR ORTHOGONAL
1. Decrease in positive evaluation	+	+	+
2. Decrease in negative evaluation	0	+	+
3. Increase in neutral "acceptance" of student initiated talk	+	+	
4. Shift from positive and negative evaluation toward neutral acknowledgment or no-evaluation of student talk by teacher	+	+	+
5. Longer periods of teacher informing	0	+	+
6. Shorter teacher questions	0	+	+
7. Increase in teacher describing	0	+	+
8. Increase in student initiated talk	+	+	+
9. Decrease in direct student responses	+		
10. Increase in divergent questions	+	+	
11. Decrease in convergent questions	+	+	
12. Decrease in evaluating interchanges			+

Key

+; score does support statement of change

0; score does not support statement of change

blank; there seems to be no score equatable

There is some indication that the observers were not always reliable and are, therefore, a source of differences. However, it should be pointed out that this lack of reliability serves to mask significant change rather than contribute to change.

Given these difficulties, it is possible to find strong substantiation for change. Overall, the intern teachers are describing more, using more divergent questions and less convergent questions, and being less evaluative and more neutral in their responses. There appears to be a shift from direct student response to student initiated responses and these student initiated responses are more often accepted or neutrally evaluated rather than praised or criticized. Further analysis will be made of the data and the results of the analysis will be reported in a formal publication.

BIBLIOGRAPHY

Amidon, B. J. and Flanders, N. A., The Role of the Teacher in the Classroom, Minneapolis: Paul Amidon and Associates, 1963.

Medley, D. M., Coding Behaviors with OScAR 4V. Office of Research and Evaluation, City University of New York: Mimeographed 1965

Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962