DOCUMENT RESUME

ED 034 741 SP 003 423

AUTHOR Medley, Donald M.; Hill, Russell A.
TITLE A Comparison of Two Techniques for Analyzing Classroom Behaviors.
PUB DATE 8 Feb 68
NOTE 16p.; Paper presented at the American Educational Research Association meeting, Chicago, Feb. 1968

EDRS PRICE EDRS Price MF-$0.25 HC-$0.90
DESCRIPTORS Behavior Patterns, *Classroom Observation Techniques, Comparative Analysis, *Interaction Process Analysis, *Measurement Instruments, *Student Behavior, *Teacher Behavior, Teacher Education, Teacher Evaluation

ABSTRACT

This document reports the results of a comparative analysis of the Flanders Interaction Analysis Technique (FIAT) and the Observation Schedule and Record, Form 4, Verbal (OScAR 4V) in terms of the kinds of information about teacher behavior which each observation technique provides. A systematic program of recorded observation (involving 70 first-year junior or senior high school teachers enrolled in an internship training program for college graduates) yielded four pairs of records made in the classrooms of each of the 70 teachers--a total of 280 pairs of records. The 10 factors considered in comparing the two techniques according to the principles of rotated factor analysis are: (1) lecturing behaviors, (2) question type, (3) question difficulty, (4) pupil initiations, (5) criticizing behavior, (6) listening behavior, (7) extended accepting behaviors, (8) question source, (9) permissive behavior, (10) managing behavior. Although each observation system gives some information not provided by the other, the authors suggest that, since OScAR 4V keys are based on twice as many basic categories as are used in FIAT, they may prove more helpful in feedback applications (i.e., in giving clearer indications of the specific kinds of behavior which a teacher might consider if he wished to change his score in a particular area). (ES)

A COMPARISON OF TWO TECHNIQUES FOR ANALYZING CLASSROOM BEHAVIORS

Donald M. Medley
Educational Testing Service

and

Russell A. Hill
Temple University

American Educational Research Association
Chicago, Illinois
February 8, 1968

A COMPARISON OF TWO TECHNIQUES FOR ANALYZING CLASSROOM BEHAVIORS

As the number of instruments available for analyzing classroom behavior grows, it becomes important to learn something about interrelationships among them. It seems worthwhile, therefore, to report the results of a study designed to compare two systems which attempt to achieve similar purposes by rather different means. Since the publication of OScAR 2a, an early instrument designed to measure classroom climate by means of live observation (Medley & Mitzel, 1958), a number of interim versions have been developed representing successive efforts to obtain useful information about other aspects of classroom climate besides the "socio-emotional" dimension defined by Withall (1949), and reflected so effectively by the "Interaction Analysis" technique of Flanders (Flanders & Amidon, 1963). This report represents the results of a study comparing Flanders' records with records obtained on a recent version of OScAR (Medley, Impellitteri, & Smith, 1966).

Procedure

The subjects of the study were 70 first-year teachers of junior or senior high school English, mathematics, science, or social studies enrolled in an internship training program for college graduates. During February each teacher was visited by a team of two observers on two different occasions; and during late May or early June each teacher was visited two more times by the same observer team. For 20 minutes on each visit, one member of each team recorded verbal behavior using Flanders' Interaction Analysis Technique (referred to as FIAT), and the other observer recorded the same behavior using the Observation Schedule and Record, Form 4, Verbal (hereinafter referred to as OScAR 4V). Twelve teachers were assigned to each team, making a total of 72, but two teachers resigned during the year, reducing the total to 70. The data of the study, therefore, consist of four pairs of records made in the classrooms of each of 70 teachers, or 280 pairs of records in all.

For those unfamiliar with the two category systems, definitions of categories in the two systems are presented in Tables 2 and 3. The observer using Flanders' system is supposed to write down the number of the category which best describes the kind of verbal behavior during each three-second interval during a visit, a FIAT record, then, consists of a string of numbers between one and ten. The task of the OScAR coder is to tally each teacher utterance (or pupil statement) in the proper cell on the special recording form included in the handout as Table 4.

The basic device used in interpreting FIAT records is a matrix of 100 cells, each containing a number representing the frequency of occurrence of a different sequence of two of the ten types of verbal behavior. A record beginning with the following sequence: 1, 2, 3, 4, 5, 6 would yield tallies in the following cells: 12, 23, 34, 45, 56, and so on. Because of the experimental dependence between successive cells--12 and 23, for instance, must share the number two--this procedure was not used in the present study. Instead, from the series 1, 2, 3, 4 5, 6 only three tallies were taken--12, 34, and 56. This would yield a matrix for each record with only half the number of tallies yielded by the normal procedure; but cell frequencies would be more nearly independent of each other.

In normal applications of FIAT, cell frequencies are converted to percents so that the total number of tallies in any matrix is 100. This procedure was not used in the present study; all scores were based on actual rather than relative cell frequencies.

Past users of FIAT have developed a number of scores called "measures," based on pooling frequencies in certain cells. Some of these measures are based on ratios of two such frequencies. In this study linear contrasts were used instead of ratios--that is, instead of a ratio of a to b the difference $a-b$ was used.

These three departures from normal scoring procedures should be borne in mind in attempts to relate findings of this study to other research with FIAT.

OScAR records were scored by combining the frequencies in the 42 cells of the recording form (Table 4) into 42 linear combinations which are all nominally orthogonal to each other. Some of the 42 scores are subtotals across groups of cells, but most of them are contrasts between groups of cells; for convenience, all of them will be called "contrasts."

Stability coefficients of the 44 FIAT measures and 42 OScAR contrasts were estimated by analysis of variance. Six FIAT measures and five OScAR contrasts had coefficients not significantly different from zero and were discarded. The remaining 75 scores were intercorrelated across all 280 records, except that between-teams variation and covariation (estimated with six degrees of freedom) was removed to prevent observer biases from distorting the correlation estimates.

It should be noted that correlations between scores on the same instrument were based on observations both made by the same observer, but that correlations across instruments were based on observations made by different observers.

Because of the practical impossibility of making sense of a 75 x 75 correlation matrix by any kind of inspection, the first 10 principal components of the matrix were extracted, and rotated to orthogonal simple structure using the varimax criterion. Interpretation was based on loadings of .50 or higher only, and will be reported here in terms of the basic categories rather than of either measures or contrasts.

Results

The ten rotated factors are listed and described in Table 1.
Of the ten, five showed some loadings of .50 or better on both instruments, and may be said to represent overlap between them. These five factors, I, II, V, VI, and VIII, accounted for 33% of the variance in the matrix. Three factors, III,

IX, and X, were unique to OScAR and accounted for another 15% of the variance. The remaining two factors, IV and VII, which accounted for 12% of the variance, were unique to FIAT. Thus, although the two techniques yield descriptions of teacher behavior which have much in common, each one also seems to get at something not readily accessible through the other system.

Each of the ten factors is described on the handout in terms of suggested factor keys which might be used to obtain scores reflecting the dimension represented by that factor. For example, from Table 1 it appears that a Factor I score could be obtained from an OScAR record by subtracting the number of Substantive Interchanges from the sum of three times the number of Continuing Informing Statements plus the number of Initiating Informing Statements tallied on the record. This factor could also be scored on FIAT by subtracting the total number of 8's (pupil responses) from the total number of 55's (steady state lecturing behaviors) on the record. This procedure would not, of course, yield exact factor scores, but should approximate them fairly well. More important, inspection of the composition of these keys is a good way to get an idea as to what behaviors enter into high or low scores on each factor.

In the case of Factor I, it appears that a high score indicates a teacher who tends not only to talk a lot, but to go on talking for a relatively long time each time he speaks, and who also tends to interact with students less than the average teacher does. In short, he lectures. Hence the factor has been tentatively named Lecturing Behaviors.

Factor II is not as easy to identify. On FIAT it is not well defined at all; inspection of the key only suggests a teacher who rejects pupil responses instead of accepting them. The key for OScAR is a complex one, reflecting the fact that the factor had substantial loadings on 11 contrasts. Inspection of the Entry totals indicates that the high scoring teacher will be one who asks

more Elaborating and Divergent Questions, and fewer Convergent ones than the average teacher. Inspection of the weights in more detail also reveals that instead of evaluating a pupil response, this teacher tends to react by asking another question which requires a pupil to elaborate on or perhaps to correct the first answer elicited. The low scoring teacher, on the other hand, tends to ask simple convergent questions and either to evaluate pupil answers as correct or merely to acknowledge them without giving any feedback information. In addition to being less likely to show any enthusiasm for pupil responses, the low scoring teacher is also less likely to praise a pupil for making a correct answer than he is to criticize him for making an incorrect one. Since this factor seems mainly to contrast teachers who prefer two opposite questioning styles--one stressing thought and more challenging, the other stressing factual knowledge and less challenging--the factor has been named Question Type.

Factor III, which does not appear on FIAT, has been named Question Difficulty. Note that scores on it are unaffected by the kind of question asked (as indicated by the zero totals on the right); this distinguishes it clearly from the dimension reflected by Factor II. Factor III seems mainly to indicate how the teacher evaluates pupil answers to whatever questions come up. The teacher scoring high on this dimension Approves pupil answers (judges them to be correct and says so to the pupils, without praise) relatively more often than he either praises them or Neutrally Rejects them--i.e., judges them to be incorrect. The last thing this teacher would do would be merely to Acknowledge a pupil's answer without indicating whether it is correct or not--particularly if the pupil is answering a Convergent question. It may be said to indicate question difficulty as reflected in the teacher's evaluations of pupils' answers.

Factor IV appears only on FIAT, and loads on a number of measures all based on category 9, Pupil Initiates. It has, therefore, been named Pupil Initiations.

This dimension seems to be defined entirely in terms of student behavior, and reflects how often a pupil speaks "because he wants to."

Factor V loads mainly on various FIAT measures based on category 7; it has therefore been named Criticizing Behavior. An inspection of the OScAR key for this factor indicates that it is based on Rebukes, but that Initiating Rebukes contribute much more to this dimension than Continuing Rebukes do. This suggests that the dimension does not reflect any deep hostility on the part of the teacher. Since it does not load on Criticizing Exits from Interchanges, Rebuking Behavior would be a better name than Criticizing Behavior.

Factor VI is the only factor defined by a single OScAR category--Continuing Pupil Statements. Its meaning is confirmed by the fact that on a FIAT record it contrasts sustained student communication--8-9 pairs--with the center of the Content Cross--4-5 pairs. Clearly this factor identifies Listening Behavior: to score high on it a teacher must do one thing; he must let pupils talk for a while without interruption.

Factor VII is the other factor found only on FIAT; since it is based on those cells in the matrix which contain only 1's, 2's, and 3's, it has been given the name Extended Accepting Behaviors, because the teacher high on the factor tends to go on accepting and praising students at length.

Factor VIII, Question Source, sounds as though it should correlate highly with Factor IV, called Pupil Initiations; but it does not. From the OScAR key it is clear that a teacher high on this factor elicits a number of Pupil-Initiated Substantive Interchanges which is high relative to the number of Teacher-Initiated ones; it is also apparent that he encourages pupils by praising them when they are right and neutrally rejecting rather than criticizing them when they are wrong. He also seems to be much more likely to acknowledge a question than to let it go unacknowledged. Only one cell in the FIAT matrix

loaded on this factor--59. Perhaps this is a clue to the way in which Factor IV is different from this factor. Only those 9's which follow a 5 reflect the behavior pattern measured by Factor VIII; Factor IV seems to reflect all pupil initiations, but Factor VIII seems to reflect only those having to do with the substantive content of the lesson.

Factors IX and X are relatively minor factors recognizable only in OScAR. Factor IX, Permissive Behavior, identifies a teacher who (1) offers pupils a choice of procedures relatively often, and (2) seldom refuses permission when asked for it. Managing Behavior consists mainly of statements which either tell pupils what to d (Directing) or discuss what they are doing, have done, or will do (Describing). The positive weight on Initiating Considering Statements probably reflects a tendency of a teacher who is discussing what a class might do to say something now and then about their feelings or desires. The negative weight on Continuing Considering Statements probably serves to filter out any genuine positive affect such utterances might contain.

Discussion

As was pointed out earlier, so far as measurement goes, the two systems appear to be rather similar, although each one gives some interesting information not provided by the other. As might have been anticipated, OScAR keys, based as they are on twice as many basic categories as are used on FIAT, provide much clearer indications of the kinds of behavior which enter into each factor than FIAT keys do. It would seem that in feedback applications this might prove a useful feature, by indicating more clearly how a teacher might proceed if he wished to change his score on that scale. The most extreme example is, perhaps, provided by Factor VIII which shows up on FIAT as a high frequency in cell 59 (teacher lectures, pupil initiates). A low score here would indicate that pupils do not voluntarily contribute to subject-matter very often, but would give no

indication of how a teacher could get them to contribute oftener. The corresponding OScAR key, on the other hand, indicates how the teacher should react to pupils' questions and to pupils' answers to teacher questions in order to score high on this dimension.

One last word. In reacting to these results one should bear in mind that they are based on da gathered in the classrooms of a rather homogeneous group of teachers, all of whom were beginners and probably constricted in their behavior for this reason. A quite different (and stronger) set of factors might be found in a different group of teachers.

References

Flanders, N., & Amidon, E. J. The role of the teacher in the classroom. Minneapolis, Minnesota: Paul S. Amidon and Associates, 1963.

Medley, D. M., Impellitteri, J. T., & Smith, Lou H. Coding teachers' verbal behavior in the classroom: A Manual for Users of OScAR 4V. New York: Division of Teacher Education of the City University of New York, 1966.

Medley, D. M., & Mitzel, H. E. A technique for measuring classroom behavior. Journal of Educational Psychology, 1958, 49, 86-92.

Withall, J. Development of a technique for the measurement of socio-emotional climate in classrooms. Journal of Experimental Education, 1949, 17, 347-361.

Table 1

Summary of Results of Joint Factor Analysis of OScAR 4V and FIAT

<table>
<tr><th></th><th>Percent of Variance</th><th>OScAR Contrast</th><th>FIAT Measure</th></tr>
<tr><td>I Lecturing</td><td>9</td><td><table><tr><th>Informing Init Cont</th><th>Substantive Interchanges</th><th>Total</th></tr><tr><td>+ 1 + 3</td><td>- 1</td><td>+ 3</td></tr></table>(Highest loading .75)</td><td>(55) vs (8)
(Highest loading .92)</td></tr>
<tr><td>II Question Type</td><td>9</td><td><table><tr><th></th><th>Sup</th><th>App</th><th>Acc</th><th>NEv</th><th>NR</th><th>CR</th><th>Total</th></tr><tr><td>CVG</td><td>+ 3</td><td>-15</td><td>-16</td><td>+ 4</td><td>- 5</td><td>- 7</td><td>-36</td></tr><tr><td>PIn</td><td>+ 2</td><td>- 2</td><td>- 4</td><td>+ 4</td><td>+ 2</td><td>- 2</td><td>0</td></tr><tr><td>DVG</td><td>- 1</td><td>+ 5</td><td>0</td><td>+ 4</td><td>+ 7</td><td>- 3</td><td>+12</td></tr><tr><td>ELB</td><td>+ 4</td><td>+ 4</td><td>+ 4</td><td>+ 4</td><td>+ 4</td><td>+ 4</td><td>+24</td></tr><tr><td>Total</td><td>+ 8</td><td>- 8</td><td>-16</td><td>+16</td><td>+ 8</td><td>- 8</td><td>0</td></tr></table>(Highest loading .81)</td><td>(86 + 87) vs (81 + 83 + 34)
(Highest loading .66)</td></tr>
<tr><td>III Question Difficulty</td><td>9</td><td><table><tr><th></th><th>Sup</th><th>App</th><th>Acc</th><th>NEv</th><th>NR</th><th>CR</th><th>Total</th></tr><tr><td>CVG</td><td>- 9</td><td>+95</td><td>-54</td><td>- 6</td><td>- 9</td><td>-17</td><td>0</td></tr><tr><td>PIn</td><td>- 6</td><td>+18</td><td>-24</td><td>0</td><td>- 6</td><td>+18</td><td>0</td></tr><tr><td>DVG</td><td>+ 3</td><td>+11</td><td>- 6</td><td>- 6</td><td>+ 3</td><td>- 5</td><td>0</td></tr><tr><td>ELB</td><td>-12</td><td>+44</td><td>-12</td><td>-12</td><td>-12</td><td>-20</td><td>0</td></tr><tr><td>Total</td><td>-24</td><td>+168</td><td>-96</td><td>0</td><td>-24</td><td>-24</td><td>0</td></tr></table>(Highest loading .87)</td><td>(No high loading)</td></tr>
<tr><td>IV Pupil Initiations</td><td>7</td><td>(No high loading)</td><td>(Highest loading .84)
Total 9</td></tr>
</table>

Table 1 (Continued)

<table>
<tr><th colspan="2"></th><th>Percent of Variance</th><th>OScAR Contrast</th><th>FIAT Measure</th></tr>
<tr><td>V</td><td>Criticizing</td><td>6</td><td><table><tr><th colspan="2">Rebuking</th><th>Total</th></tr><tr><th>Init</th><th>Cont</th><th></th></tr><tr><td>+ 3</td><td>+ 1</td><td>+ 4</td></tr></table>(Highest loading .68)</td><td>Total 7
(Highest loading .89)</td></tr>
<tr><td>VI</td><td>Listening Behavior</td><td>5</td><td>Continuing Pupil Statements
(Highest loading .85)</td><td>(88 + 89 + 98 + 99) - (44 + 45 + 54 + 55)
(Highest loading .56)</td></tr>
<tr><td>VII</td><td>Indirect Teacher Talk</td><td>5</td><td>(No high loading)</td><td>(Highest loading .86)
(11 + 12 + 13
+21 + 22 + 23
+31 + 32 + 33)</td></tr>
<tr><td>VIII</td><td>Question Source</td><td>4</td><td><table><tr><th></th><th>Sup</th><th>App</th><th>Acc</th><th>NEv</th><th>NR</th><th>CR</th><th>Total</th></tr><tr><td>PIn</td><td>+ 1</td><td>0</td><td>+ 4</td><td>0</td><td>+ 1</td><td>0</td><td>+ 6</td></tr><tr><td>TIn</td><td>- 1</td><td>0</td><td>- 4</td><td>0</td><td>- 1</td><td>0</td><td>- 6</td></tr><tr><td>Total</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td><td>0</td></tr></table>(Highest loading .66)</td><td>(59)
(Highest loading .60)</td></tr>
<tr><td>IX</td><td>Permissive Behavior</td><td>3</td><td><table><tr><th colspan="4">Non-Substantive Question</th></tr><tr><th colspan="2">Teacher</th><th colspan="2">Pupil</th><th rowspan="2">Total</th></tr><tr><th>Pos</th><th>Neg</th><th>Pos</th><th>Neg</th></tr><tr><td>+ 2</td><td>0</td><td>0</td><td>- 2</td><td>0</td></tr></table>(Highest loading .84)</td><td>(No high loading)</td></tr>
<tr><td>X</td><td>Managing Behavior</td><td>3</td><td><table><tr><th></th><th>DSC</th><th>DRC</th><th>CNS</th><th>Total</th></tr><tr><td>Int</td><td>+ 2</td><td>+ 2</td><td>+ 1</td><td>+ 5</td></tr><tr><td>Cont</td><td>+ 2</td><td>+ 2</td><td>- 1</td><td>+ 3</td></tr><tr><td>Total</td><td>+ 4</td><td>+ 4</td><td>0</td><td>+ 8</td></tr></table>(Highest loading .65)</td><td>(No high loading)</td></tr>
</table>

Table 2

Summary of Categories of Verbal Behavior on OScAR 4V

I. STATEMENTS

A. Teacher Statements--utterances which neither respond to nor solicit a response from a pupil--are classified as follows:

1. AFFECTIVE. A statement revealing sensitivity to pupil feelings is classified as CONSIDERING. A statement criticizing pupil conduct is classified as REBUKING.

2. SUBSTANTIVE. A statement containing no affect but referring directly to content to be learned by pupils is classified as INFORMING if it conveys a fact, generalization, or the like, or PROBLEM STRUCTURING if it sets up a question or issue to be solved.

3. PROCEDURAL. A statement which contains neither affect nor substance is classified as DIRECTIVE if it contains a command or instruction with the force of a command. A statement which does not clearly fall into one of the above categories is classified as DESCRIBING.

B. Pupil Statements--utterances by pupils addressed to other pupils are classified as PUPIL STATEMENTS.

C. Sequence. If a teacher makes two or more successive statements which may be classified in the same category, all except the first are classified as CONTINUING. The first statement in a series of the same kind is classified as INITIATING.

II. INTERCHANGES

An interchange is an episode in which a pupil says something to the teacher and the teacher reacts.

A. Substantive Interchanges are those in which the pupil's utterance refers to content to be learned. Such interchanges contain two parts: entry and exit.

1. Entries. A substantive interchange begins with one of four types of entries:

a. PUPIL INITIATED. The pupil addresses a statement or question to the teacher.

b. ELABORATING. The teacher addresses a question to a pupil which refers directly to a previous pupil comment.

Table 2 (continued)

c. DIVERGENT. The teacher addresses a question to a pupil which does not refer directly to a previous pupil comment, and which offers him a choice of two or more acceptable or "correct" answers.

d. CONVERGENT. The teacher addresses a question to a pupil which does not refer directly to a previous pupil comment and to which there is only one acceptable answer.

2. Exits from Completed Substantive Interchanges. After the pupil has asked his question or made his answer, the teacher disposes of the answer in one of six ways, called Exits. Exits are first classified according to the information they contain about the correctness or acceptability of what the pupil has said.

If the teacher clearly indicates that what the pupil has said is correct or acceptable, the interchange is classified as SUPPORTED if praise or enthusiasm is shown, as APPROVED if praise is not given.

If the teacher clearly indicates that what the pupil has said is incorrect or unacceptable, the interchange is classified as CRITICIZED if disapproval of either the pupil or what he has said is expressed, or as NEUTRALLY REJECTED if no disapproval is expressed.

If the teacher makes some response to what the pupil says which does not clearly indicate whether it is correct (acceptable) or incorrect (unacceptable), the interchange is classified as ACCEPTED; if the teacher makes no response, it is classified as NOT EVALUATED.

B. Non-Substantive Interchanges are those in which the pupil's contribution does not refer to content to be learned.

1. TEACHER-INITIATED non-substantive interchanges are classified as POSITIVE or NEGATIVE according to the affective content of the teacher's question.

2. PUPIL-INITIATED non-substantive interchanges are classified as POSITIVE if the teacher supports, approves, or accepts the pupil's suggestion, and as NEGATIVE if he criticizes, neutrally rejects, or ignores it.

Table 3

Summary of Categories for Interaction Analysis

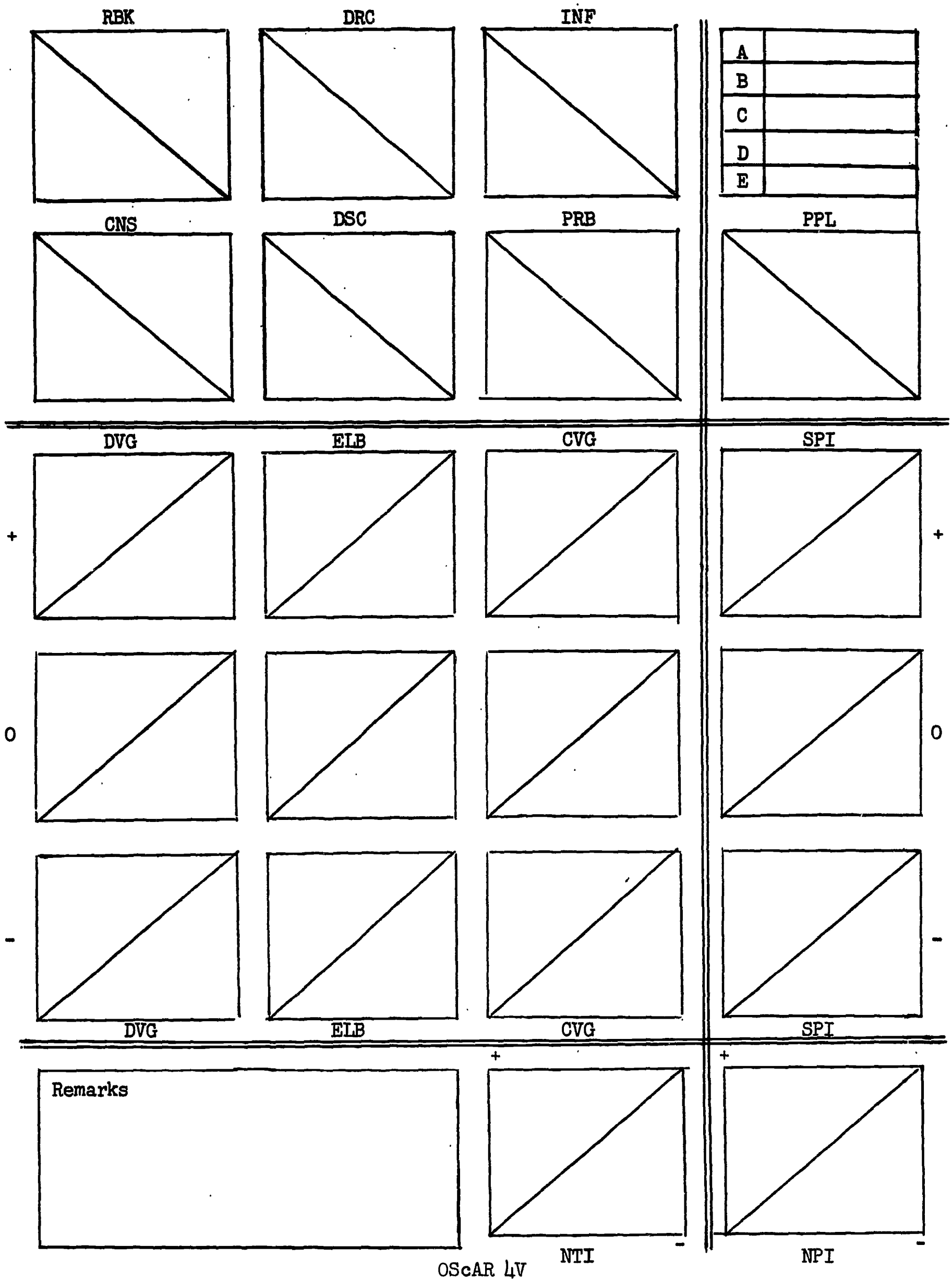
Teacher Talk	Indirect Influence	1.* Accepts Feeling: accepts and clarifies the feeling tone of the students in a non-threatening manner. Feelings may be positive or negative. Predicting or recalling feelings are included.
		2.* Praises or Encourages: praises or encourages student action or behavior. Jokes that release tension, not at the expense of another individual, nodding head or saying "um hm?" or "go on" are included.
		3.* Accepts or Uses Ideas of Student: clarifying, building, or developing ideas or suggestions by a student. As teacher brings more of his ideas into play, shift to category five.
		4.* Asks Questions: asking a question about content or procedure with the intent that a student answer.
	Direct Influence	5.* Lecturing: giving facts or opinions about content or procedure; expressing his own ideas, asking rhetorical questions.
		6.* Giving Directions: directions, commands, or orders to which a student is expected to comply.
		7.* Criticizing or Justifying Authority: statements intended to change student behavior from non-acceptable to acceptable pattern; bawling someone out: stating why the teacher is doing what he is doing; extreme self-reference.
Student Talk		8.* Student Talk--Response: talk by students in response to teacher. Teacher initiates the contact or solicits student statement.
		9.* Student Talk--Initiation: talk by students which they initiate. If "calling on" student is only to indicate who may talk next, observer must decide whether student wanted to talk. If he did, use this category.
		10.* Silence or Confusion: pauses, short periods of silence and periods of confusion in which communication cannot be understood by the observer.

*No scale is implied by these numbers.

Adapted from:

Flanders, N. A. Teacher influence, pupil attitudes, and achievement. Minneapolis: University of Minnesota (U. S. Office of Education Cooperative Research Project No. 397), 1960. (Mimeographed)

Table 4



OScAR 4V