DOCUMENT RESUME

ED 065 550 TM 001 695

AUTHOR	Limbacher, Philip C.; Rosenshine, Barak
TITLE	Relationship of High-Inference and Low-Inference Observation Measures.
PUB DATE	Apr 72
NOTE	13p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, Illinois, April 1972)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Correlation; Effective Teaching; *Interaction Process Analysis; Lesson Observation Criteria; Performance Criteria; *Rating Scales; Statistical Analysis; *Student Attitudes; Tables (Data); *Teacher Evaluation
IDENTIFIERS	IA; Interaction Analysis; Teacher Performance Appraisal Scale; TPAS

ABSTRACT

The correlation between high-inference measures and low-inference observation measures was studied using two groups of social studies student teachers (25 students per group). The high-inference measure for two video-taped lessons was obtained with the 10-item Teacher Performance Appraisal Scale (TPAS), on which pupils rated the lesson on a scale from one to seven. The low-inference observation measures were obtained by having trained observers categorize the videotapes into the 10 categories of Classroom Interaction Analysis (IA) observation system of Ned A. Flanders. Separate analyses were made for each group. No steady correlations were established between the high-inference and low-inference data. Another study was then conducted in which scores for the individual items comprising the TPAS were developed, and the results were subjected to additional correlation studies with the IA scores. The only consistent correlation was a negative correlation across observations for the high-inference variable "content organized" and the low-inference variable "student talk--initiation." (DB)

RELATIONSHIP OF HIGH-INFERENCE AND LOW-INFERENCE OBSERVATION MEASURES

Philip C. Limbacher
Texas A&M University

Barak Rosenshine
University of Illinois

Paper presented at the annual meeting of the American Educational Research Association, April 3 - 7, 1972, Chicago, Illinois

RELATIONSHIP OF HIGH-INFERENCE AND LOW-INFERENCE OBSERVATION MEASURES

A search to establish correlations between high-inference data, such as those obtained through fairly specific questionnaires, and lower-inference data, such as those obtained through the use of observational category systems, is useful for several reasons. Many of the variables discussed in the literature that have consistently related to student achievement were secured by the use of high-inference rating scales, e.g., clarity, variability, enthusiasm, etc.[1] However, the lists of specific teacher behaviors describing these variables have not been developed. If, within those high-inference variables found to correlate most consistently with student achievement, certain specific teacher behaviors can be isolated, the task of teaching these skills to pre-service and in-service teachers becomes considerably easier. At present many colleges and universities do teach their beginning teachers specific teacher behaviors, i.e., questioning skills, reinforcement techniques, etc. The justification for the selection of these behaviors appears at present to be based mainly on expert belief and seldom on experimental or correlational studies linking the specific behavior to student achievement. The rising trend to use "performance criteria" in the training of teachers indicates that more schools will orient their programs toward specific teacher behaviors. A knowledge of which teacher behaviors relate to demonstrated effective teacher characteristics should enable curriculum developers to select specific behaviors for inclusion in their programs with more precision and confidence.

The data for this study were obtained from a previous study of two groups of social studies student teachers.[2] Each group contained twenty-five students and was observed twice, during the first week and during the sixth or last week of student teaching. The design of the study called for a thirty-minute videotape to be made of each observed lesson followed by a pupil evaluation of that

lesson. The high-inference measure for each lesson was obtained with the ten-item Teacher Performance Appraisal Scale (TPAS) developed by Horace E. Aubertine and William D. Johnson. These high-inference measures were determined by asking the pupils to rate the lesson using a scale from one to seven. The variables upon which these measures were obtained can be summarized as follows:

1. Learning aims understood
2. Learning aims developed
3. Content meaningful
4. Content organized
5. Method used appropriate
6. Method stimulating
7. Method successful
8. Contact with learner
9. Participation encouraged
10. Feeling of accomplishment

The last ten minutes of each period observed were reserved for use by the investigator who read a definition of each of the high-inference measures and attempted to describe the results a pupil should have seen for various levels of each measure. These uniform definitions of the individual items attempted to standardize the information given the pupils by the several investigators involved in the study.

The low-inference observation measures were obtained by having trained observers categorize the videotapes into the ten categories of the classroom interaction analysis (IA) observation system developed by Ned A. Flanders.[3] The observers were trained using the instructions and ground rules contained in the manual developed by Edmund J. Amidon and Ned A. Flanders.[4] The ten categories used in this system can be summarized as follows:

1. Accepts feelings
2. Praises or encourages
3. Accepts or uses student ideas
4. Asks questions
5. Lecturing
6. Giving directions
7. Criticizing or justifying authority
8. Student talk--response
9. Student talk--initiation
10. Silence or confusion

In this system the observer determines at three-second intervals which category is being used. The resulting tallies were processed by computer program to develop totals, percentages, and selected ratios for each lesson.

Ned Flanders, the developer of this low-inference observation system, feels that the use of these measures can help to explain why the results were obtained in studies on teaching effectiveness and, in fact, lists as one of the two questions his latest book on teaching behavior attempts to speak to as, "Can objective data provide insights into subjective problems?" His discussion of this concept concludes with this statement, "It may take years of research and development before we can synthesize the subjective and objective elements, but there can be no escape from confronting the question."[5] We have attempted to confront the question; and our results have been confusing, disappointing, and suggestive for future studies.

Because the two groups of social studies student teachers differed in the training they had prior to student teaching, separate analyses were made for each group. We hoped to find significant relationships between the high-inference and IA variables which would be consistent across the two groups and across the two occasions. Table I indicates the results, showing any significant correlations between one of the Flanders IA categories and the high-inference variables. Thus, looking at the table, you can assume that there were no significant correlations at all between Flanders IA categories 1, 4, 6, 7, and 9. The TPAS instrument has been factor analyzed by William D. Johnson, University of Illinois, who developed a TPAS-A cognitive oriented factor and a TPAS-B affective oriented factor from the TPAS variables. These two factors were also used in this analysis.

The original study secured observations on another high-inference measure, the Illinois Teacher Evaluation Questionnaire (ITEQ), developed by the late Richard E. Spencer, University of Illinois. This instrument was administered

only once at the completion of student teaching. The ITEQ secured a measure of pupil evaluation of total teacher performance, e.g., tests, homework, textbooks, etc. The significant correlations for this measure and the four ITEQ sub-factors created by Spencer are also included in Table I.

Our search to establish correlations between the high-inference measures of the TPAS and ITEQ and the lower-inference measures of Flanders IA received little support from this initial analysis. An examination of Table I reveals that there is no correlation between a Flanders behavior and a TPAS which holds up across trials and across groups. The correlations that do exist would appear to have limited generality.

In an attempt to include only those student teachers for whom we had complete measure on all three scales, the number of teachers used was reduced to a total of seventeen from both groups. As the ITEQ measure of total teacher performance again gave only scattered and inconclusive correlations and was administered only at the conclusion of the student teaching experience, our attention was focused upon the two measures available for both observations--Flanders and TPAS. Table II indicates the results of this analysis. And, again, we were unable to get, for this kind of population, steady correlations between the Flanders variables and high-inference measures. Correlations were developed for several IA categories for the first observation, but none of these held up across observations. A comparison of the mean scores for each of the TPAS and IA variables across the two occasions explains some of the problem. Tables III and IV show the means and standard deviations of the twenty variables across trials.

Observation indicates that teacher scores on the high-inference measure were slightly lower on the second administration. To the extent that pupil evaluation of these lessons is also a measure of attitude toward the student

teacher, this decrease in positive attitude is a fairly universal phenomenon and might be expected in any longitudinal study. In contrast, the variables on the Flanders category system moved both up and down; indeed, when an entire category system is considered, the scale is ipsative. Thus, as we now realize, the scores on the student rating form (TPAS) were moving down across time, whereas, the scores on the category system were moving in an ipsative manner. Given these phenomena, one might expect that in other studies in which both ratings and specific counts were obtained on the same teachers at two points in time, the correlations would not be consistent across occasions.

An examination of Table IV shows that consideration might be given to the use of other measures of central tendency and deviations for data containing categories with such extreme measurements. The median could be expected to yield a more representative value, particularly in the case of categories six and nine which indicate sigmas greater than the mean for both observations. The use of a quartile deviation could be combined with the median to provide more representative values.

Table V was developed to determine the correlation for each of the IA categories across the two observations. The stability across observations of the three main teacher categories used by the student teachers seems to be revealed by this table.

Unable to establish steady correlations between our high-inference and low-inference data, we developed scores for the individual items comprising the high-inference measure (TPAS) and subjected the results to further correlational studies with the IA low-inference scores. The entire population of student teachers available for both observations, forty for the first and thirty-two for the second, was subjected to a missing-data intercorrelation analysis. The results of this analysis can be found in Tables VI and VII. We again were unable

to secure consistent correlations between the variables with but one exception. This time we were able to achieve a negative correlation across observations for the high-inference variable "Content organized" and the low-inference variable "Student talk--initiation." When the high-inference measure was examined as a single entity or as one of two factors, this phenomenon remained obscured. It would seem that, at least for this population of junior and senior high school pupils, when the student teacher conducts a lesson in which the pupils are free to initiate verbal behavior, these same students will subsequently judge that lesson low on the organization of its content. The material read to the pupils as they were asked to evaluate the variable "Content organized" of the lesson stated, "If the teacher makes little or no effort to differentiate between relevant or irrelevant topics, the lesson is not organized. If you cannot logically order the major points of the lesson into a pattern meaningful to you, the lesson was not organized."[6] In other words, if the pattern or organization of the lesson was not clear to the pupil, the lesson deserved a low rating. The cognitive clarity or perceived organization of a teacher's lesson is one of the few high-inference variables found to correlate with regression-adjusted student achievement scores.[7] The search for low-inference behaviors which comprise "organization" would seem to have revealed a specific classroom behavior, the quantity of self-initiated student talk, which produces a negative effect on the perceived level of this important high-inference variable. Educators, attempting to train their students to present lessons most likely to achieve high ratings on clarity or organization and using methods designed to elicit or encourage self-initiated student talk, should be aware of this finding.

In conclusion, the difficulties which we encountered in attempting to relate high-inference variables to low-inference variables when data are collected at two points in time may cast doubt upon the results obtained when data are collected

only once and may suggest the fallibility of attempts to relate ratings of teachers by students to counts of specific behavior.

---

[1]Barak Rosenshine and N. Furst, "Research in Teacher Performance Criteria," in Research in Teacher Education, ed. by B. Othanel Smith (New Jersey: Prentice-Hall, Inc., 1971).

[2]Philip C. Limbacher, "A Study of the Effects of Microteaching Experiences Upon Practice Teaching Classroom Behavior" (unpublished doctoral dissertation, University of Illinois, 1968).

[3]Ned A. Flanders, Teacher Influence, Pupil Attitudes and Achievement, Final Report Cooperative Research Project No. 397, United States Office of Education (Minneapolis: University of Minnesota, 1960).

[4]Edmund Amidon and Ned A. Flanders, The Role of the Teacher in the Classroom (Minneapolis: Paul S. Amidon and Associates, 1963).

[5]Flanders, Analyzing Teaching Behavior (Massachusetts: Addison Wesley Publishing Company, 1970).

[6]William D. Johnson, "A Discussion of Individual Items of the Teacher Performance Appraisal Scale" (University of Illinois, 1967, Mimeographed).

[7]Rosenshine and Furst, "Research in Teacher Performance Criteria," pp. 44-45.

TABLE 1

SIGNIFICANT CORRELATIONS BETWEEN FLANDERS IA CATEGORIES AND HIGH-INFERENCE MEASURES (TPAS and ITEQ)

	TRIAL 1			TRIAL 2		
	N = 20 EXPERIMENTAL	N = 20 CONTROL	N = 40 BOTH	N = 13 EXPERIMENTAL	N = 13 CONTROL	N = 26 BOTH
F2 & TPAS	.59**	.00	.19	.15	.36	.24
F2 & TPAS-A	.63**	.01	.22	.12	.39	.25
F2 & TPAS-B	.49	-.02	.13	.18	.27	.20
F3 & TPAS	.03	.44	.27	.04	.57*	.22
F3 & TPAS-A	.07	.46*	.29	.08	.53	.23
F3 & TPAS-B	-.06	.37	.21	-.06	.56*	.18
F5 & TPAS	-.53*	-.38	-.31*	-.24	-.16	-.09
F5 & TPAS-A	-.53*	-.38	-.30	-.06	-.10	.01
F5 & TPAS-B	-.52*	-.37	-.31	-.54	-.24	-.25
F8 & TPAS	.42	.46*	.36*	.17	.46	.15
F8 & TPAS-A	.43	.52*	.39*	.15	.41	.14
F8 & TPAS-B	.39	.33	.29	.17	.49	.16
F10 & TPAS	.13	.21	.16	-.17	-.58*	-.36
F10 & TPAS-A	.14	.25	.19	-.32	-.56*	-.42*
F10 & TPAS-B	.11	.11	.10	.13	-.55	-.21
F2 & ITEQ				-.31	-.49	-.42*
F2 & method				-.31	-.46	-.39*
F2 & item				-.14	-.55*	-.42*
F2 & knowledge				-.59*	-.49	-.52**
F2 & interest				-.32	-.39	-.34

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

TABLE II

SIGNIFICANT CORRELATIONS BETWEEN FLANDERS IA AND TPAS FOR SELECTED SUBJECTS (N = 17)

	First Observation			Second Observation		
Flanders	TPAS-Total	TPAS-A	TPAS-B	TPAS-Total	TPAS-A	TPAS-B
1	.00	.00	.00	.00	.00	.00
2	.49*	.53*	.41	.20	.19	.19
3	.30	.34	.22	.26	.30	.15
4	.67**	.73**	.54	.03	.04	.00
5	-.68**	-.69**	-.63**	-.12	.00	-.35
6	-.08	-.12	-.01	-.08	-.21	.17
7	.16	.15	.18	.46	.43	.45
8	.56*	.59*	.47	.09	.09	.08
9	.02	-.03	.13	.14	.02	.36
10	-.21	-.21	-.20	-.22	-.33	.05

** $p \leq .01$
* $p \leq .05$

TABLE III

MEANS AND SIGMAS OF HIGH-INFERENCE VARIABLES ACROSS OBSERVATIONS (TPAS)

	First Observation (N = 40)		Second Observation (N = 32)	
High-Inference Variable	Mean	Sigma	Mean	Sigma
Learning aims understood	4.93	.41	5.04	.46
Learning aims developed	4.87	.46	4.85	.49
Content meaningful	4.99	.59	5.10	.59
Content organized	5.27	.47	4.98	.56
Method appropriate	5.16	.52	5.15	.52
Method stimulating	4.58	.89	4.68	.76
Method successful	5.14	.49	5.02	.54
Contact with learner	5.40	.57	5.32	.62
Encourages learner participation	4.94	.56	4.85	.60
Feeling of accomplishment	4.91	.66	4.85	.61

TABLE IV

MEANS AND SIGMAS OF LOW-INFERENCE VARIABLES ACROSS OBSERVATIONS (IA)

Low-Inference Variable	First Observation Mean(%)	Sigma	(N-using Category)	Second Observation Mean(%)	Sigma	(N-using Category)
Accepts feelings	.00	.00	0	.00	.00	0
Praises or encourages	2.47	1.81	28	1.32	1.10	21
Accepts student ideas	3.74	3.13	31	3.60	3.16	25
Asks questions	14.30	6.55	32	12.96	6.63	26
Lecturing	34.19	20.10	32	33.96	22.96	26
Giving directions	3.7	5.53	23	2.87	4.63	16
Criticizing	1.4	1.31	13	.73	.70	13
Student talk--response	21.24	14.20	32	24.50	15.91	26
Student talk--initiation	14.9	19.10	29	12.71	15.76	23
Silence or confusion	7.76	5.50	32	10.70	8.12	26

TABLE V

CORRELATIONS OF LOW-INFERENCE VARIABLES ACROSS OBSERVATIONS

Low-Inference Variable	Correlation Coeff. (Trial I with Trial II)	N-using Category
Accepts feelings	.00	0
Praises or encourages	.34	18
Accepts student ideas	.67**	24
Asks questions	.57**	26
Lecturing	.48*	26
Giving directions	.15	11
Criticizing	-.30	7
Student talk--response	.01	26
Student talk--initiation	.23	20
Silence or confusion	.31	26

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

TABLE VI

CORRELATIONS BETWEEN INDIVIDUAL ITEMS FOR FLANDERS IA AND TPAS
FIRST OBSERVATION

	N for each variable	Learning Aims Understood	Learning Aims Developed	Content Meaningful	Content Organized	Method Appropriate	Method Stimulating	Method Successful	Contact with Learner	Encourages Learner Participation	Feeling of Accomplishment
	40	40	40	40	40	40	40	40	40	40	40
Accepts feelings	0	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Praises or encourages	27	.07	.04	-.01	-.04	-.05	-.07	.02	-.15	.12	.19
Accepts student ideas	30	.18	.31	.08	.38*	.25	.12	.21	.21	.22	.39*
Asks questions	31	.24	.24	.16	.07	.09	.08	.15	.01	.18	.16
Lectures	31	-.38*	-.14	-.40*	.11	-.35*	-.37*	-.37*	-.01	-.40*	-.26
Gives directions	22	-.11	-.24	-.42*	-.04	-.21	-.16	-.14	-.32	-.28	-.09
Criticizes	13	-.16	.06	-.12	.01	-.13	.15	.08	.11	-.09	.12
Student talk--response	31	.37*	.37*	.37*	.22	.32	.21	.40*	.29	.35*	.39*
Student talk--initiation	29	.02	-.22	.14	-.37*	.01	.12	-.07	-.16	.07	-.19
Silence or confusion	31	.16	.04	.08	.03	.27	.24	.33	-.06	-.01	.18

* $p < .05$

TABLE VII

CORRELATIONS BETWEEN INDIVIDUAL ITEMS FOR FLANDERS IA AND TPAS
SECOND OBSERVATION

	N for each variable	Learning Aims Understood	Learning Aims Developed	Content Meaningful	Content Organized	Method Appropriate	Method Stimulating	Method Successful	Contact with Learner	Encourages Learner Participation	Feeling of Accomplishment
	32	32	32	32	32	32	32	32	32	32	32
Accepts feelings	0	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Praises or encourages	21	.04	.14	.13	.21	.13	.14	.31	.04	-.07	.13
Accepts student ideas	25	.21	.04	.02	.33	.09	.07	.15	.11	.04	.10
Asks questions	26	.19	.14	.07	.32	-.04	.00	.19	.32	.12	.25
Lectures	26	-.12	.23	-.08	.40*	-.13	-.32	-.11	-.03	-.24	-.04
Gives directions	16	-.31	-.18	-.17	-.19	-.07	-.12	-.23	.13	.10	-.06
Criticizes	13	.31	-.18	.06	-.20	.12	.09	-.03*	-.24	.16	-.01
Student talk-- response	26	.16	.07	.07	.06	.07	.17	.24	.14	.08	.24
Student talk-- initiation	23	.03	-.28	.16	-.56**	.30	.31	-.03	-.17	.21	-.14
Silence or confusion	26	-.34	-.49**	-.32	-.55**	-.13	-.15	-.41**	-.31	-.14	-.38*

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