DOCUMENT RESUME

ED 104 959 TM 004 419

AUTHOR Swartz, Ned K.
TITLE Divergent Perception of Teaching Effectiveness by Different Groups of Raters.
PUB DATE 1 Apr 75
NOTE 17p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Washington, D. C., April 1, 1975); Best copy available
AVAILABLE FROM Dr. Ned Swartz, School of Education, Old Dominion University, Norfolk, Virginia 23508

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Administrator Attitudes; Effective Teaching; *Evaluation Methods; High School Students; High School Supervisors; Peer Groups; Rating Scales; Secondary Education; Self Evaluation; Statistical Analysis; *Teacher Evaluation; *Teacher Rating; *Trade and Industrial Teachers; Vocational Education Teachers

ABSTRACT

A mail survey was conducted to study the differences among group ratings of the same instructor. A population of 72 teachers of differing background levels, teaching in the field of trade and industrial education, was selected. Each instructor was rated by one school administrator, a local supervisor, two teaching colleagues, all of the students enrolled in one of the classes of the instructor, and a self evaluation. Four different evaluation instruments were developed for the study; 15 items were common to all four instruments and these were used in the analysis. Significant differences were found in the performance rating evaluations of all four groups. Statistical data from the study is included. (DEP)

DIVERGENT PERCEPTION OF TEACHING EFFECTIVENESS BY DIFFERENT GROUPS OF RATERS

Ned K. Swartz
Old Dominion University
P. S. Vivekananthan
Virginia Polytechnic Institute and State University

A paper presented at the annual meeting of the National Council of Measurement in Education, Washington, D. C., April 1, 1975. Requests for copies should be sent to Dr. Ned Swartz, School of Education, Old Dominion University, Norfolk, Virginia 23508.

DIVERGENT PERCEPTION OF TEACHING EFFECTIVENESS BY DIFFERENT GROUPS OF RATERS

Ned K. Swartz
Old Dominion University
P. S. Vivekananthan
Virginia Polytechnic Institute and State University

INTRODUCTION

Evaluation of instructors is a customary process in any educational institution. A proclaimed justification of the evaluation procedure is to serve administrative purposes (i.e. promotion, tenure and salary adjustment) and to help teachers to improve the instructional process. The controversy of who should evaluate instructors is widespread among researchers.

More frequently than not, students are asked to evaluate their teachers. A major criticism levied against student evaluation of teaching performance is that it becomes a popularity contest among teachers and, as a result, a teacher who is popular among students, whether he is good at teaching or not, would be rated high. Smyser (1948:459), Cole (1940:569) and more recently Brickman (1966:143) have expressed blunt opposition to the use of student evaluations. On the other hand, Guthrie (1954:1-21) and McKeachie, et.al. (1971:437) have argued favorably for the use of student evaluations of teaching performance.

Teacher superiors are another source of information for the teacher's classroom behavior. Superiors often hold administrative positions and, thus, are criticized on the grounds that they do not have direct classroom contact with the instructor and can not effectively evaluate classroom behavior. Hedlund (1953:231-232) and Johnson and Radebaugh (1969:152-156), on the other hand, found that administrators could effectively evaluate and identify superior teachers.

A third source of teaching effectiveness is the teacher-peer. Theoretically, it appears sound that teacher-peers are in a good position to evaluate adequately because they are not only in constant contact with their fellow instructors but also have students in common who can informally feed them information about the instructor's classroom behavior. However, Morsh and Wilder (1954) found that instructors tended to evaluate their colleagues upon the amount of subject matter which they possessed rather than upon their actual teaching effectiveness.

The fourth method of teacher evaluation is ratings by the teachers themselves. Very little research relating to the use of teacher self-ratings has been done. Studies which have been conducted resulted in inconclusive results about their value. Turner (1971:1-98) found that teachers generally rate themselves lower than principals or students on certain performance traits such as class management, professional attitudes and growth, and personal qualities. Bolton (1973:140-141) favored self-evaluation with the contention that it provides an instructor the opportunity to improve his teaching skills by observing his own behavior in a threat-free atmosphere.

Considering the results of these studies, the question is: Would these groups rate differently when asked to rate the same instructor? A related question is: If there are differences in the group ratings, what are the dimensions separating the groups? The present study addresses these questions.

PROCEDURE

Subjects

From a total population of 453 trade and industrial education instructors in Virginia, a stratified sample of 72 teachers with different background levels were selected for the study. Each teacher was to be rated by (a) one school administrator (director, principal, assistant principal, or assistant director), (b) one

supervisor (local school division director or supervisor of vocational education, general supervisor, or secondary supervisor), (c) two teacher-peers in the subject area of vocational industrial education, (d) one self-rating and (e) all students enrolled in one class taught by the instructor. Responses were obtained from 67 school administrators (group 1), 71 supervisors (Group 2), 134 teacher-peers (Group 3), 68 self-ratings (Group 4), and 796 students (Group 5). All surveys were conducted by mail.

Instruments

Four evaluation instruments were developed for the study. One instrument was used by Group 1 and Group 2, one by Group 3, one by Group 4, and the fourth by Group 5. Expert opinions of state and national leaders in the field of vocational industrial education were used in the design of item content of the instruments. A pilot study showed that that test - retest reliability of the instruments ranged from .70 to .95 for the five groups. There were 15 items common to all four instruments and were subsequently used in the analysis.

Stepwise discriminant analysis was used to identify those dimensions which separated the five groups of raters.

RESULTS

Items common to all rating forms used in the discriminant analysis are shown in Table 1.

The means and standard deviations, by groups, for each item are presented in Table 2.

The discriminant analysis yielded Wilks' Λ of .79298 with 15 and 41131 d. f. The approximate F value was 4.45 with 60 and 4362.43 degrees of freedom, which is significant at the .05 level. As a result, it was concluded that there were significant differences in the mean group performance ratings. Further analysis

showed that there were no differences among Groups 1, 2 and 3 while Groups 4 and 5 differed not only between themselves, but also from the other groups. The results of this analysis are shown in Table 3.

Thus, the five groups formed three clusters, one cluster consisting of administrators, division supervisors and teacher-peers, the second cluster consisting of teachers, and the third consisting of students.

Discriminant analysis also computed the sum of the squares on dimension differences, namely eigen values. The root lambdas (λ's), or the eigen values, are $\lambda_1 = .17487$, $\lambda_2 = .04413$, $\lambda_3 = 0.1800$, $\lambda_4 = .00983$, and $\lambda_i = .000$, where i = 5, 6. . .15. The relative sizes of λ indicate the extent to which the associated discriminant functions distinguish between groups. The Rao's chi-square approximation (1952:373) was used to test the significance of each discriminating function (λ). The results of this test are shown in Table 4.

As a result of this test, the first and second roots were found to be significant beyond the .005 level, the third is significant beyond .025. These λ's accounted for 96 per cent of the total variance. Therefore, it was concluded that there were three dimensions separating the five groups.

The group centroids for each dimension were calculated and are shown in Table 5.

The locations of each of the group centroids, in two dimensional space are illustrated by the graphs in Figures 1, 2, and 3.

The evidence of three clusters is very apparent when the group centroids are plotted on the first two dimensions - which accounts for 88.61 percent of the total variance.

Identification of Dimensions

The coefficients for canonical variables show the degree to which each item on the rating scale contributes to each dimension. In order to have a basis for identification of each dimension, only those items which had a coefficient of .35 or greater were used. Items relevant to dimension I are given in Table 6.

An analysis of the items identified on dimension I reveals that item 1, "Professionalism" - Personal Characteristics, measures personal attributes of the instructor which influence learning. Item 2, Course Content, deals with the relevancy of the course, the use of a trade advisory committee, and clear progression from each unit of the course toward the final goal of preparation for work. Items 9, Communication Skills, and 10, Teaching Methods, provide insight into the instructor's actual teaching methods, types of instructional activities, and ability to communicate with students. Item 15, Overall Performance, provides an analysis of the total instructional ability of the teacher. Item 1 seems to be unrelated to the other items; however, it has a degree of relationship in that personal characteristics of an instructor may have some influence on the learning of students. Items 8, 9, 10, and 15 each relate to the manner in which an instructor attempts to convey information. This dimension, therefore, is best described as "Teaching Proficiency," since this concept deals with the skills and competencies used in the education process.

Table 7 describes items of dimension II. On dimension II, item 3, Motivation of Students, relates to the way the instructor instills in his students a desire to learn. Item 4, Classroom Control, is a measure of the instructor's ability to maintain discipline in the classroom or laboratory. Items 5, 6, and 14 which deal with Vocational Subject Knowledge, Teacher Interest in the Subject, and Safety and Safety Instruction are concerned with the ability of the instructor to provide

meaningful instructional skills to his students. They also deal with the knowledge and interest he shows in the trade subject he is teaching. As a result, dimension II was labeled, "Instructor's Ability to Promote Learning" because each item appeared to relate in some way to this central concept.

Dimension III related to the adjustments necessary on the part of the vocational industrial education instructor to his professional surroundings. The items are given in Table 8. Items 2 and 6, "Professionalism" - Staff Relations and Teacher Interest in Subject, measure his relationship with others on the professional staff and his interest in the subject he is teaching. His interest in the student organization in vocational industrial education and the appreciation and care which he displays for the physical facilities and equipment used in teaching are measured in item 12, Support of Student Organizations, and item 13, Use and Care of Physical Facilities and Equipment. Items 5 and 8, Vocational Subject Knowledge and Course Content, are concerned with the teacher's ability to adjust to these surroundings by his willingness to improve his knowledge and his "Insturctional Plan" of imparting trade skills to students. Consequently, dimension III was given the title "Relationship to Educational Surroundings."

CONCLUSIONS AND IMPLICATIONS

There are significant differences in the performance evaluation ratings given by school administrators, school division supervisors, teacher-peers, teachers (self-rating) and students. It implies that the emphasis of teaching effectiveness is placed differently by different groups and, thus, ratings from a single group of raters would not reveal a total picture of the teaching effectiveness of an instructor. The ratings of school administrators, school division supervisors, and teacher-peers are similar. This result concurs with that of Owens (1971) who concluded that administrators, teachers, and college supervisors perceive most areas of teacher competence similarly. Probably these three groups are familar with

the teacher's background with regard to such factors as educational level, teaching experience and trade experience, while students may not have this information about the teacher. As a result, students rated differently. In any case, it is suggested that in order to get an adequate picture of an instructor's teaching performance, evaluations must be obtained from several groups who are familar with his classroom behavior.

Table 1

Items Common to All Rating Forms Used in Discriminant Analysis

Item Number	Item
1	"Professionalism" - Personal Characteristics
2	"Professionalism" - Staff Relations
3	Motivation of Students
4	Classroom Control
5	Vocational Subject Knowledge
6	Teacher Interest in Subject
7	Organization
8	Course Content
9	Communication Skills
10	Teaching Methods
11	Evaluation Techniques
12	Support of Student Organizations-VICA*
13	Use and Care of Physical Facilities and Equipment
14	Safety and Safety Instruction
15	Overall Performance

*Vocational Industrial Clubs of America

Table 2

Group Means and Standard Deviations for Each Item

Variable	Administrators		Supervisors		Teacher-Peers		Self-Rating		Student	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
1	4.19403	0.82091	4.04225	0.86910	4.17910	0.87445	4.01471	0.70165	4.15829	0.80408
2	4.23881	0.85439	4.07042	0.86724	4.28358	0.82818	4.19118	0.75818	4.29774	0.81986
3	3.92537	0.65859	3.78873	0.79130	3.94030	0.83855	4.04412	0.65640	4.13065	0.87764
4	4.07463	0.82229	4.02817	0.71657	4.02985	0.85787	4.13235	0.71036	4.13693	0.98926
5	4.25373	0.57617	4.39437	0.72661	4.45522	0.77183	4.20588	0.65923	4.55276	0.76294
6	4.25373	0.76561	4.04225	0.76413	4.15672	0.79334	4.39706	0.62628	4.45729	0.79582
7	3.86567	0.85094	3.80282	0.78594	3.93284	0.84246	3.73529	0.74548	4.00879	0.86938
8	3.58209	1.03205	3.57746	0.83942	3.83582	0.96708	3.47059	0.87196	4.19849	0.83973
9	3.74627	0.78515	3.90141	0.77745	4.01493	0.81328	3.73529	0.80330	4.16457	0.89821
10	3.86567	0.69403	3.70422	0.86840	3.94030	0.83855	3.75000	0.83532	3.79271	1.01976
11	3.85075	0.72307	3.84507	0.72993	3.94776	0.78823	3.91176	0.76753	4.08794	0.93283
12	3.62687	1.12594	3.71831	1.08474	3.62687	1.23648	3.35294	1.21881	4.04146	1.04102
13	4.20895	0.80786	4.18310	0.76176	4.15672	0.86584	4.04412	0.81833	4.43593	0.79597
14	4.14925	0.72307	4.14084	0.72300	4.15672	0.78380	3.86765	0.80861	4.36181	0.81785
15	4.04478	0.68386	3.92958	0.72356	4.17910	0.74441	3.95588	0.63325	4.32035	0.82817

Table 3

F-Matrix Comparison of Group Means

Group Number	Administrators (1)	Supervisors (2)	Teacher-Peers (3)	Self-Rating (4)
Supervisors (2)	0.86470			
Teacher-Peers (3)	1.53299	1.14612		
Self-Rating (4)	2.25166*	2.62607*	3.42016*	
Student (5)	5.39985*	3.84348*	5.04472*	7.14566*

*Significant at .05 level

Table 4

Signifigance of the Discriminant Function
Chi-Square χ^2 Approximations

Function	λ	d.f.	χ^2	p
1	.17487	18	176.782*	<.05
2	.04413	16	43.914*	<.05
3	.01800	14	22.520*	<.05
4	.00983	12	10.130	>.05

Table 5

Group Centroids for Each Dimension

Rater Group (Group Number)	Dimension		
	I	II	III
School Administrators (1)	.80360	.20205	-.26868
Supervisors (2)	.56165	.25461	-.22158
Teacher Peers (3)	.43864	.25648	.29533
Teachers on a Self Rating (4)	.86081	-.70009	.05739
Students (5)	-.26512	-.02308	-.01224

Table 6

Dimension I. "Teaching Proficiency"

Item Number	Item	Coefficient
1	"Professionalism" - Personal Characteristics	0.43340
8	Course Content	-0.89946
9	Communication Skills	-0.39620
10	Teaching Methods	0.50674
15	Overall Performance	-0.40136

Table 7

Dimension II. "Instructor's Ability to Promote Learning"

Item Number	Item	Coefficient
3	Motivation of Students	-0.59308
4	Classroom Control	-0.35435
5	Vocational Subject Knowledge	C.[illegible]288
6	Teacher Interest in Subject	-1.12833
14	Safety and Safety Instruction	0.41932

Table 8

Dimension III. "Relationship to Educational Surroundings"

Item Number	Item	Coefficient
2	"Professionalism" - Staff Relations	0.49367
5	Vocational Subject Knowledge	-0.60828
6	Teacher Interest in Subject	-0.42508
8	Course Content	0.49596
12	Support of Student Organizations	-0.50692
13	Use and Care of Physical Facilities and Equipment	-0.47472
15	Overall Performance	0.82198

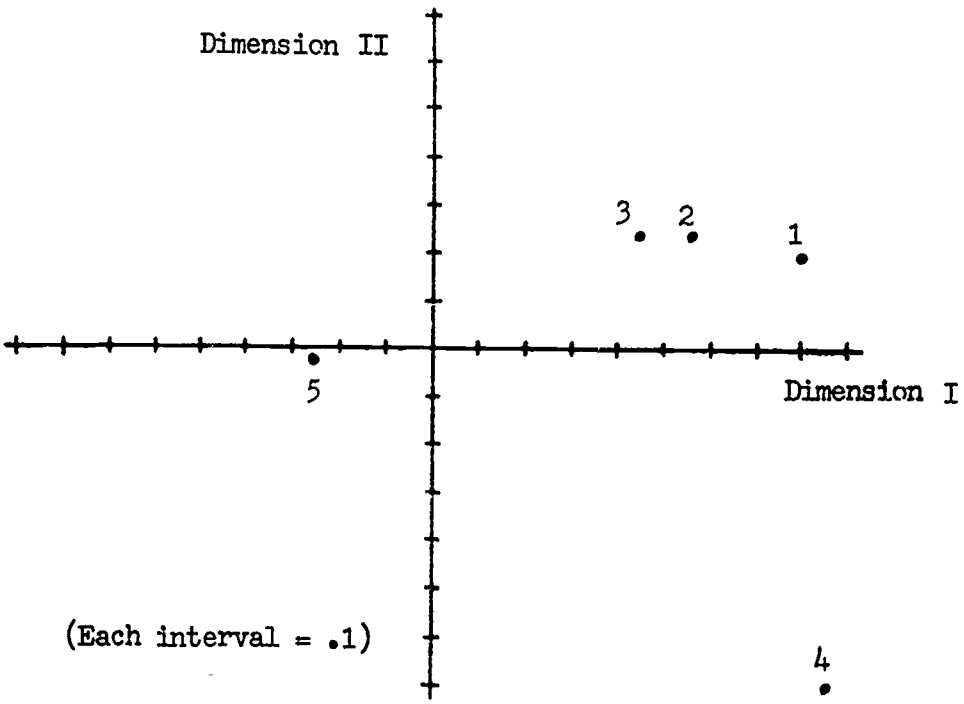

Figure 1

Plot of Centroids for Dimensions I & II

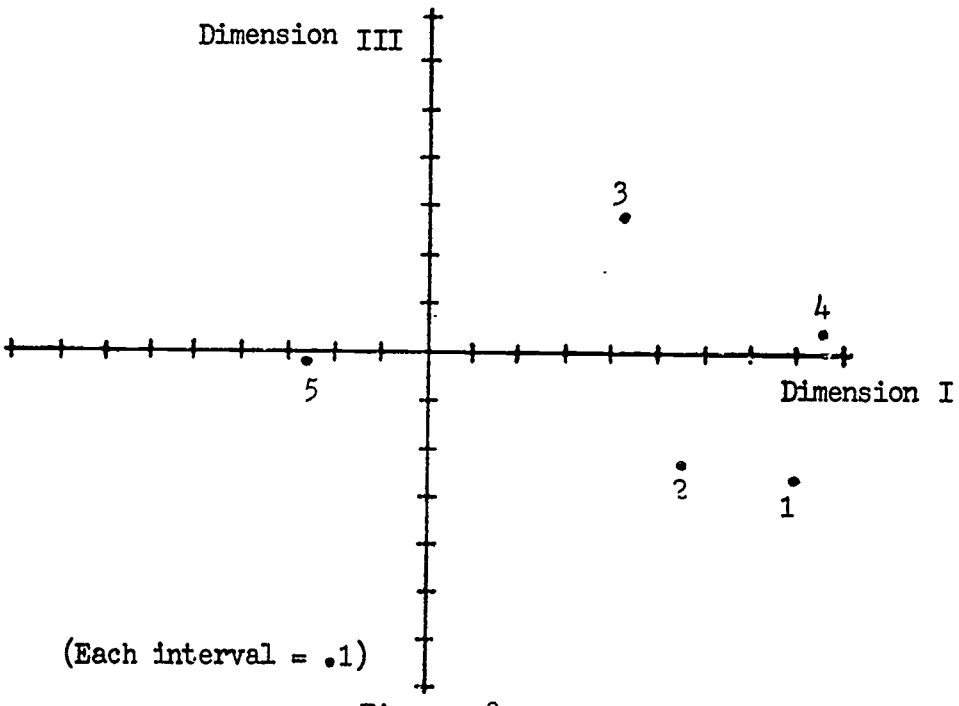

Figure 2

Plot of Centroids for Dimensions I & III

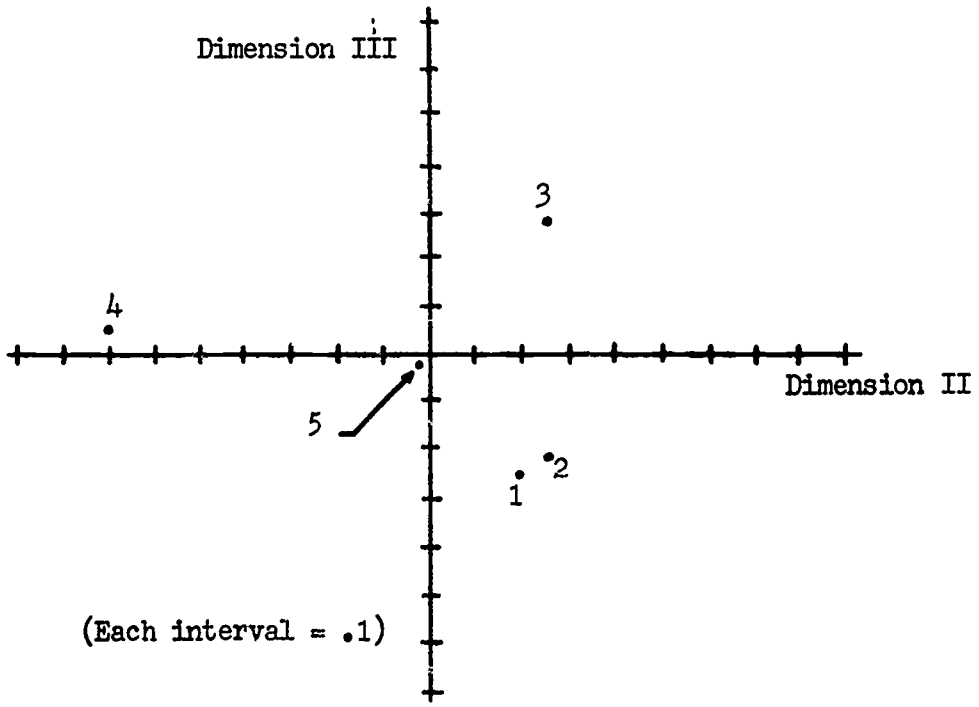

Figure 3

Plot of Centroids for Dimensions II & III

LITERATURE CITED

Bolton, Dale L. Selection and Evaluation of Teachers. Berkley, California: McCruchan Publishing Corp., 1973

Brickman, W.W. "Student Evaluation of Professors," School and Society, 94 (1966), 143.

Cole, Luella. The Background for College Teaching. New York: Farrarr and Rinehart, 1940.

_____. The Evaluation of Teaching--Progress Report. Seattle: University of Washington, 1954

Hedlund, Paul A. "Cooperative Study to Predict Effectiveness in Secondary School Teaching," The Journal of Teacher Education, XLIV (September, 1953) 231-232.

Johnson, James A., and Byron F. Radebaugh. "Excellent Teachers--What Makes Them Outstanding?" The Clearing House, XLIV (November, 1969), 152-156.

McKeachie, William J., and others. "Student Ratings of Teacher Effectiveness: Validity Studies," American Educational Research Journal, 8 (1971), 437.

Morsh, Joseph H., and Eleanor W. Wilder. Identifying the Effective Instructor: A Review of the Quantitative Studies 1900-1952. Lackland Air Force Base, Texas: Training Aids Research Laboratory, Air Force Personnel and Training Research Center, Research Bulletin AFPT RC-TR-54.

Smyser, W. C. "Whom the Shoe Fits. . . ," College and University, 23 (April, 1948) 459.

Turner, Max L. "Teaching Effectiveness as Perceived by Pre-Tenure Teachers, Their Students and Principals." Unpublished Ed. D. dissertation, Utah State University, 1964.