American Educational Research Journal
Summer 1985, Vol. 22, No. 2, Pp. 267–272

Dimensions Underlying Student Ratings of Instruction: A Multidimensional Scaling Analysis

MARTHA L. BANZ and JOSEPH LEE RODGERS
University of Oklahoma

Multidimensional scaling (MDS) was used to derive the dimensions underlying student ratings of instruction. A weighted MDS revealed that departmental affiliation, course structure, and intensity of teaching were particularly salient to the raters. The findings are discussed in light of the current literature on student ratings of instruction.

Considerable research effort has been spent considering the psychometric, theoretic, and logistic problems involved in the construction and administration of instruments for student ratings of instruction. Of particular concern has been the question, "What are the teacher characteristics that students use in making their evaluative ratings?" A number of relevant characteristics have been cited in the literature, including organization, rapport, skill, course structure, dynamism, fairness, enthusiasm, course difficulty, and use of humor (see Centra, 1979; Cohen, 1981; Dowell & Neal, 1982; Marsh, 1980). In addition, a number of methodological techniques have been used in the assessment of this question, including factor analysis, analysis of variance, and several multivariate techniques. These methods have measured the characteristics assumed to underlie the ratings; the important dimensions are defined by the researcher. While the results of these studies are highly suggestive, they are always open to the criticism that some important dimension of teaching behavior was overlooked and not measured.

A relatively new analytical model—the multidimensional scaling (MDS) model—can be used to address this issue. Several characteristics make this technique particularly attractive for this purpose. First, the model has not been previously used in this assessment. Second, the relevant dimensions underlying teaching behavior are derived from the data, rather than being

Downloaded from http://aerj.aera.net at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016
Downloaded from SHARE.OK. Persistent Handle URL: https://hdl.handle.net/11244/25563

imposed on the subjects in the data collection process. Finally, MDS results can be used in conjunction with regression analysis to link previous techniques that have been used to the approach presented here.

MDS refers to a class of statistical procedures that transforms proximity information into a spatial configuration of points reflecting the underlying structure of a data set. The methods for converting the proximities into points in a Euclidean space are based on the theorems of Young and Householder (1938) and subsequent work by Torgerson (1958). In the present context, MDS allows the derivation of the dimensions underlying student ratings.

METHOD

Subjects

Fifty-eight undergraduate students from a small liberal arts college in Oklahoma served as subjects. All were enrolled in upper division courses in the social and behavioral sciences. Ninety-three percent indicated that they knew at least five of the eight instructors in the Division.

Instruments

The first instrument used, a similarities instrument, consisted of a list of the 28 possible pairs of instructors from the Division. The pairs were listed in random order and counterbalanced so that no instructor was consistently listed first within a pair. Following each pair was a 4½-inch line labeled "exact same" at one end and "very different" at the other. The similarity judgments given on this instrument were used to derive the MDS configuration.

A second set of instruments was used to rate each instructor on communication skills, fairness, intensity of teaching, student relations, course structure, influence on students, enthusiasm, course workload, interest as a lecturer, appearance, course difficulty, dynamism/charisma, personal lifestyle, and classroom atmosphere. These scales were derived primarily from the research literature. Following each scale was a 4½-inch line spanning the continuum for each dimension, with explanatory labels at each end. These responses were used to validate the MDS configuration.

Procedure

Subjects were given the similarities instrument with a cover sheet indicating that they were to consider the instructors' teaching qualities in the classroom. Subjects were instructed to mark an X on the line to indicate the perceived similarity between each pair of instructors. Upon completing the first form, the subject returned it and was given the rating forms, where an X was marked along the line indicating the rating for each instructor on each of the 14 scales.

Downloaded from http://aerj.aera.net at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

Numerical values of the similarity judgments were derived by dividing the response line into nine equal units and observing where the X was marked. This produced a square, symmetric matrix of similarities between the eight instructors for each subject. Numerical values for the rating scales were derived by the same method. An additional nominal scale of departmental affiliation was defined by assigning instructors in the same department the same numerical value, producing a total of 15 scales. This yielded an instructor by rating scale rectangular matrix for each of the subjects.

To account for individual differences among subjects, a weighted MDS analysis was performed. ALSCAL (Young & Lewyckyj, 1979) was used to analyze the similarity judgments. Analyses were performed in two, three, and four dimensions using ordinal and interval measurement levels.

RESULTS

Following the guidelines given in Kruskal and Wish (1978), the three-dimensional, ordinal solution was chosen for interpretation. This solution was chosen on the basis of goodness-of-fit and interpretability. The stimulus configuration is given in Figure 1. In this figure, each letter reflects one instructor and the distance between the points reflects the degree of dissimilarity between instructors.

The three dimensions of the space were interpreted as intensity, course structure, and departmental affiliation. The vertical dimension corresponded to intensity. In general, Instructors H, B, F, and E are less intense in their teaching style than are A, C, D, and G. The horizontal dimension was interpreted as course structure. Instructors D and H teach loosely structured courses, while Instructor C's are highly structured. The last dimension corresponded almost perfectly to departmental affiliation. Instructors A, B, F, and H are behavioral scientists (psychologists, sociolo-

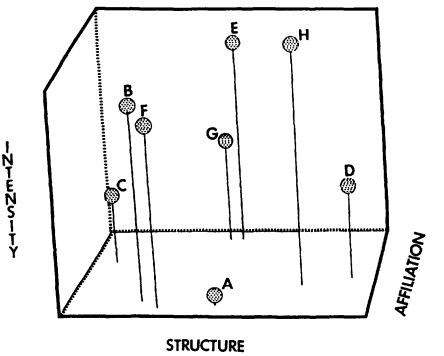

FIGURE 1. Stimulus (instructors) space for three-dimensional, ordinal MDS configuration.

Downloaded from http://aerj.aera.net at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

gists), whereas G, E, and C are social scientists (political scientists, historians).

The interpretations presented above were based on the researchers' knowledge of these instructors, and, therefore were somewhat subjective. To validate these dimensions, a linear regression analysis was performed using the 15 rating scales. The regression analysis, with the stimulus coordinates as the predictors and the rating scales as criteria, defined the least-squares projection of the rating scales into the stimulus space. The resultant squared multiple correlation (R^2) indicated the variability in the rating scale accounted for by the stimulus configuration.

Five of the rating scales had R^2 values that exceeded .90 and were statistically significant ($p < .05$). In descending order, these were difficulty, structure, affiliation, intensity, and workload. Inspection of the direction cosines suggested the three axes corresponded to departmental affiliation, intensity of teaching, and course structure. Table I provides information concerning some characteristics of the instructors, along with the R^2 values for the relevant dimensions.

Two other rating scales, difficulty and workload, had high R^2 values and were highly correlated. In weighted MDS, however, the interpretable directions are expected to closely correspond to the axes (Kruskal & Wish, 1978). According to the direction cosines, the best interpretations for the axes through the stimulus space were affiliation, intensity, and structure.

TABLE I

Mean Ratings and R^2 Values for Instructors Across 14 Rating Scales

Rating scale	Instructor								R^2
	A	B	C	D	E	F	G	H	
Course difficulty	6.7	5.7	7.5	5.2	5.9	5.8	6.5	4.8	.98***
Course structure	7.2	6.7	8.0	5.7	6.0	6.7	7.0	5.7	.96**
Intensity of teaching	6.2	5.0	6.3	5.0	5.4	4.8	5.8	4.3	.95**
Course workload	6.5	5.3	7.2	5.0	5.6	5.7	6.0	4.9	.90*
Personal lifestyle	3.8	4.7	5.0	4.0	5.1	4.4	4.9	3.5	.82
Classroom atmosphere	4.2	2.9	3.9	3.2	4.2	3.4	3.9	2.7	.67
Fairness in grading	7.1	7.6	7.5	7.3	6.6	7.5	7.0	7.3	.63
Enthusiasm	7.3	7.7	8.2	7.3	6.5	6.9	6.7	6.6	.56
Interest as lecturer	5.2	5.0	8.5	6.7	6.0	5.8	6.7	4.9	.48
Dynamism/charisma	5.7	6.5	7.7	6.6	5.9	5.4	6.8	5.5	.48
Communication skills	6.3	7.7	8.1	7.1	6.3	7.1	7.3	6.1	.43
Personal appearance	6.8	7.4	7.4	8.0	6.8	6.8	8.0	7.6	.26
Influence on students	6.7	7.5	7.9	7.3	6.4	6.6	6.9	7.1	.13
Relations with students	6.7	7.8	7.8	7.3	6.5	6.9	7.1	7.6	.09

Note. Departmental affiliation has been omitted from this table to protect the anonymity of the instructors being rated. Its R^2 was .95.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Downloaded from http://aerj.aera.net at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

It is interesting to note the negligible relationships of factors such as appearance, influence on students, and student relations (all $R^2 < .26$). Also, interest as a lecturer, enthusiasm, dynamism, and communication skills had only moderate relationships ($.43 < R^2 < .56$), suggesting that these characteristics are not as salient as is often supposed.

In addition to the group stimulus configuration, weighted MDS also produces a subject space. Inspection of this space indicated that most subjects used all three dimensions in making their similarity judgments, with substantial variability in the students' judgments being accounted for by the derived dimensions.

DISCUSSION

Many statistical techniques have been used to define and assess student ratings of instruction. Previous research indicates that such ratings are multidimensional with a number of factors proposed as relevant to the instructional process. This study identified which of those many dimensions were most salient to the student rater. Using MDS, the underlying dimensions of student ratings were derived rather than imposed. These were departmental affiliation, intensity of teaching, and course structure. Difficulty and workload were also relevant, though not as interpretable.

Two measurement tasks were used in this study. The similarity task is fundamentally different from that used in most student evaluations. The traditional approach to evaluation corresponds more closely to the second task in this study, that is, rating the instructor on a number of proposed scales. Asking the student to make similarity judgments between instructors raises the possibility of their using any number of objective criteria, such as age, sex, rank, and so on. It is interesting that only one of the three derived dimensions is objective (affiliation) while the other two are relatively subjective (structure and intensity). Uncovering two dimensions that are subjective lends support to an assumption implicit throughout this study, that the dimensions that are salient to raters in making similarity judgments are also salient to them in making the more traditional evaluative ratings.

These findings support many of the results found in the literature on student ratings. Unique to this investigation, however, is the perceived salience of departmental affiliation. Although affiliation is not necessarily related to good teaching, when students are asked to make *similarity* judgments about instructors' teaching behavior, the academic discipline of the instructor is salient. Affiliation might have an indirect influence on ratings in that the preference of a student for any given discipline may influence their predisposition to like or dislike the instructor teaching that subject.

Certainly, the dimensions of teaching that are most salient to the student

Downloaded from http://aerj.aera.net at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016

rater are not the only ones of interest. Nevertheless, knowledge of the aspects of instruction to which students pay particular attention can help in the interpretation of the evaluation instruments commonly used. It is hoped that this new perspective can give insight into the complex relationship between instructional evaluation and teaching effectiveness.

REFERENCES

CENTRA, J. A. (1979). *Determining faculty effectiveness.* San Francisco: Jossey-Bass.

COHEN, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Review of Educational Research, 51,* 281–309.

DOWELL, D. A., & NEAL, J. A. (1982). A selective review of the validity of student ratings of teaching. *Journal of Higher Education, 53,* 51–62.

KRUSKAL, J. B., & WISH, M. (1978). *Multidimensional scaling.* Beverly Hills, CA: Sage Publications.

MARSH, H. W. (1980). The influence of student, course, and instructor characteristics in evaluations of university teaching. *American Educational Research Journal, 14,* 441–447.

TORGERSON, W. S. (1958). *Theory and methods of scaling.* New York: John Wiley.

YOUNG, F. L., & LEWYCKYJ, R. (1979, October). The ALSCAL procedure (SAS Tech. Rep. S-113). Chapel Hill, NC: University of North Carolina.

YOUNG, G., & HOUSEHOLDER, A. S. (1938). Discussion of a set of points in terms of their mutual distances. *Psychometrika, 3,* 19–22.

AUTHORS

MARTHA L. BANZ, Graduate Student, Department of Psychology, University of Oklahoma, Norman, OK 73019. *Specializations:* Quantitative methods, educational psychology.

JOSEPH LEE RODGERS, Assistant Professor of Psychology, Department of Psychology, University of Oklahoma, Norman, OK 73019. *Specializations:* Measurement and evaluation, behavioral statistics.

Downloaded from http://aerj.aera.net at UNIV OF OKLAHOMA LIBRARIES on January 20, 2016