Journal of Applied Psychology
1995, Vol. 80, No. 4, 491–499

0021-9010/95/$3.00

Negativity (and Positivity) in Performance Evaluation: Three Field Studies

Yoav Ganzach
Tel-Aviv University

Because of a lack of powerful nonlinear models, there is little research about nonlinearity in performance evaluation in nonexperimental, real-world data. Nonlinearity in 3 real-world data sets of performance evaluation was examined by using various versions of a nonlinear model labeled the *scatter model.* The findings indicate that performance evaluations tend to be conjunctive, that is, more weight is given to negative attributes than to positive attributes. However, this basic tendency disappears when the overall level of evaluation is high, as a result of inconsistency resolution—the tendency to resolve inconsistency between 2 or more aspects of the input information on the basis of an overall evaluation.

There is much interest in the literature in the subject of nonlinearity in performance evaluation and, in particular, in the use of configural integration rules—rules in which the impact of a given cue depends on its value relative to the values of other cues (e.g., Fusilier & Hitt, 1983; Hitt & Barr, 1989; Jago, 1978; Stumpf & London, 1981). Two types of configural rules are of major interest in the domain of performance evaluation, the disjunctive rule and the conjunctive rule. A *disjunctive rule* is a rule in which rating is based primarily on positive attributes, and a *conjunctive rule* is a rule in which rating is based primarily on negative attributes. (The terms *negativity bias* and *positivity bias* are often used to describe these rules. See Skowronski & Carlston, 1989, and Kanouse & Hanson, 1972, for a review.) Indeed, there are a number of studies that have examined whether raters pay more attention to positive or negative information in performance evaluation. These studies almost unanimously suggest that the negative information receives more attention (Bolster & Springbett, 1961; Brannick & Brannick, 1989; Hamilton & Huffman, 1971; Hollman, 1972; London & Hakel, 1974; London & Poplawski, 1976; Webster, 1964; Wyer & Hinkle, 1976; see DeNisi, Cafferty, & Meglino, 1984, for a review).

All previous studies that examined for nonlinearity in performance evaluation, and for disjunction–conjunction in particular, were laboratory studies. In those studies, the cues on which evaluations were based were usually nonrepresentative in that intercue correlations were zero (e.g., Ogilvie & Schmitt, 1979; Weldon & Mustari, 1988), the setting was usually artificial, and participants were aware that they were involved in an experiment. Therefore, in this article, I investigated nonlinearity in performance evaluation on the basis of field data. This investigation is important for two reasons. First, it offers practitioners simple procedures for understanding the processes underlying performance evaluation in their organizations on the basis of routinely collected internal data, without resorting to complex experimental procedures. Second, it overcomes the problem of external validity, which is viewed by many researchers as a serious drawback in laboratory research about performance evaluation (e.g., Bernadin & Vilanova, 1986; Campbell, 1986; Guion, 1991; Guion & Gibson, 1988; Ilgen & Favero, 1985).

What is the reason for the lack of research about reliance on nonlinear strategies outside of the laboratory? In my view, the reason is the lack of powerful nonlinear models that are appropriate for studying nonexperimental field data. Field data are usually characterized (once attributes are rescaled to have a positive correlation with the criterion) by positive intercorrelations between attributes. When intercorrelations are positive, it is difficult to discriminate among different models in general, and among linear and nonlinear models in particular. In this case, the predictions from the different models are highly correlated, and the power of tests aimed at discriminating between the models is quite low (Einhorn, Kleinmuntz, & Kleinmuntz, 1979).

The Davidson and the K-Mart centers supplied financial support for this research.

I thank Avraham Kluger and Dan Mingelgrin for their comments on earlier versions of this article and Lea Alon and Hana Keidar for helpful discussions.

Correspondence concerning this article should be addressed to Yoav Ganzach, School of Business Administration, Tel-Aviv University, Tel-Aviv 69102, Israel.

Recently, however, a nonlinear model that has the potential of detecting nonlinearity in field data was introduced by Brannick and Brannick (1989) and was further developed by Ganzach (1993, 1994b; Ganzach & Czaczkes, 1995). Therefore, in the current research this model, labeled the *scatter model,* was used to investigate nonlinearity in various nonexperimental data sets that were obtained in field settings.

The article is organized as follows: First, I describe the data sets that were analyzed. Second, I analyzed these data sets by using two versions of the scatter model—the simple version and the multiple-scatter version. Finally, I used these two versions to analyze some effects associated with the relationship between nonlinear strategies and inconsistency resolution, namely, the tendency to resolve inconsistency between two or more aspects of the input information on the basis of an overall evaluation (Ganzach, 1994a).

Data Sets

Employee Evaluations

Psychologists evaluated 448 candidates for a technical job on four attributes (cognitive skills, technical skills, working style, and personality) using a 9-point scale with 1 representing a very positive evaluation and 9 representing a very negative evaluation. The attribute scores were determined on the basis of interviews and standard tests. In addition, the psychologists provided a general evaluation of the candidate on a similar 9-point scale.

Student Evaluations

Applicants for the medical school at the Hebrew University of Jerusalem are interviewed by a committee of three faculty members. Each faculty member evaluates the candidate according to 13 attributes on a 10-point scale with 1 representing a very positive evaluation and 10 representing a very negative evaluation (the attributes pertain primarily to personality characteristics such as persistence, consistency, sociability, morality, maturity, and motivation, or to behavioral characteristics in the interview, such as tension, reaction to authority, emotional tone, and intellectual responsiveness) and subsequently provides a general evaluation on a 5-point scale with 1 representing a very positive evaluation and 5 representing a very negative evaluation. The data set contains 1,577 evaluations of 638 candidates (observations with missing values, either on the attributes or on the general evaluation, were excluded from the data set).

Class Evaluations

This data set contains students' class evaluations at the Hebrew University of Jerusalem in the spring semester of 1993. The data set contains 14,905 evaluations of 2,168 classes (observations with missing values were excluded). The evaluations are completed at the end of each semester in each of the classes taught at the Hebrew University of Jerusalem. The results of these evaluations are used, among other things, for faculty promotion decisions, so great care is taken by the university in the collection and the processing of the data. The form used to obtain the evaluations is standard. Ten questions on the form pertain to the evaluation of specific class attributes, that is, to the extent to which (a) "the course developed [the student's] reasoning and analytical abilities," (b) "the tutorial-laboratory added to the understanding of the lectures," (c) "the homework added to the course," (d) "the course was planned and organized," (e) "the readings added to [the student's] understanding of the course material," (f) "the lectures were interesting," (g) "the teacher lectured in a clear and organized manner," (h) "the professor achieved the right mix between lecturing and student participation," (i) "the professor answered and responded to comments and criticism," and (j) "the professor treated the students properly." In addition, there are two general evaluation questions, one pertaining to the course ("What is your general evaluation of the course?") and one pertaining to the professor ("What is your general evaluation of the professor?"). All evaluations are expressed on a 20-point scale with 1 representing a very negative evaluation and 20 representing a very positive evaluation.

The Simple Version of the Scatter Model

Overview

Consider an overall evaluation of two job candidates on the basis of two equally important attributes. The two candidates have the same mean score; however, one has two moderate scores and the other has one high score and one low score. If judgments follow a linear rule, the two candidates would receive a similar evaluation. However, if judgments follow nonlinear, or configural, rules, the evaluations of the two candidates may be different. If they follow a disjunctive rule, the high-scatter candidate, the candidate characterized by high discrepancy between attribute values, would receive a higher score. If they follow a conjunctive rule, the low-scatter candidate would receive a higher score.

The two-candidates example illustrates the relationship between scatter and prototypical nonlinear strategies in a two-attribute case. When the number of attributes is larger, reliance on such nonlinear strategies can be estimated by various versions of a model that takes into account the scatter(s) among the various attributes. The simplest version of this model is given by

$$Y = \alpha + \sum_{i=1}^{m} \beta_i X_i + \delta \,\mathrm{SXT}, \qquad (1)$$

where Y is the overall evaluation, X_is are the evaluations of the various attributes, and SXT is the scatter, defined as

$$\left[\sum_{i=1}^{m}(X_i-\bar{X})^2\right]^{1/2},$$

where $\bar{X}$ is the mean X_i within each profile,

$$\bar{X}=\sum_{i=1}^{m}X_i/m.$$

The X_is are standardized and rescaled (if necessary) to have a positive correlation with Y.

This model represents judgment by two elements: the elevation of each profile—a weighted average of the attribute values—and the scatter of the profile—the standard deviation of the (standardized) attribute values around the profile mean (see Cronbach & Gleser, 1953, for treatment of the concepts of elevation and scatter). The influence of the profile's scatter on judgment is indicative of reliance on disjunctive or conjunctive rules. If a disjunctive rule is used, scatter will be positively related to judgment, whereas if a conjunctive rule is used, scatter will be negatively related to judgment.

Analyses

On the basis of Equation 1, four regressions were estimated, one for each of the four evaluation tasks (class evaluations included two evaluation tasks: evaluation of the course and evaluation of the professor). Prior to the analysis (as well as the other analyses reported in this article), Blom's transformation (Blom, 1958; Tukey, 1962) was performed on the data to reduce skewness. Such a transformation is recommended in the study of the performance of nonlinear models (Goldberg, 1976). However, the analyses reported below were performed on the untransformed data, and the statistical conclusions were similar.

For all four tasks, the scatter coefficients were significantly negative ($p < .0001$), which indicates reliance on conjunctive strategy. Thus, the results of the analyses of these field data are consistent with previous laboratory experiments that showed that people assign higher weights to negative information than to positive information in performance evaluation.

To supply the reader with an idea of the magnitude of the effects, I calculated the partial correlation between scatter and rating, controlling for elevation (attributes' values). The partial correlations of each of the rating tasks are given in column 2 of Table 1. (The results of these analyses are obviously similar to the results of the analyses based on the scatter coefficient.)

In previous research, nonlinear models were evaluated by comparing their cross-validated fit to the cross-validated fit of the linear model (e.g., Einhorn, 1970, 1971; Goldberg, 1971). The cross-validated multiple correlations squared of the linear model and the simple version of the scatter model appear in Table 2. (These are the averages of the two cross-validated multiple correlations squared that resulted from dividing each data set by an odd–even split and using each half once as a modeling sample and once as a holdout sample.) It is clear from these results that the cross-validated multiple correlations squared of the scatter model are higher than those of the linear model. Although the difference between the two sets of multiple correlations squared is not large, it is important to note that even small differences in model fit may be associated with large differences in underlying strategies. It is well-known that because the linear model is quite robust to deviations from linearity, it gives a good fit even to judgments that are generated by nonlinear strategies (Dawes & Corrigan, 1974; Johnson, Meyer, & Ghose, 1989).

The Multiple-Scatter Version

Overview

The simple version of the scatter model does not take into account the structure of the attribute information and, in particular, the dimensional structure of this information. No knowledge about the nature of the judgment task is needed to examine for configurality; a simple cookbook approach is used in which scatter information is aggregated and represented by a single general scatter term.

However, when attributes are (perceived to be) organized in dimensions, the configural impact of an attribute belonging to a certain dimension may be determined more by its position relative to attributes belonging to that dimension than by its position relative to attributes belonging to other dimensions. Thus, configurality may be associated more with intradimension scatters than with a general scatter. In addition, except for intradimension scatters, configurality may also be associ-

Table 1
Partial Correlations Between Rating and Scatter

	Partial correlation		
Task	Entire data set	High half	Low half
Employee evaluation	−.175**	−.067	−.256**
Student evaluation	−.162**	−.075	−.176**
Professor evaluation	−.090**	.136**	−.256**
Course evaluation	−.096**	.016	−.194**

** $p < .0001$.

Table 2
Cross-Validated Fit of the Linear Model and the Simple Version of the Scatter Model

	Cross-validated R^2	
Task	Linear	Scatter
Employee evaluation	.859	.863
Student evaluation	.604	.617
Professor evaluation	.818	.820
Course evaluation	.739	.741

ated with interdimension scatter, which represents the scatter between the dimensions. For example, Ganzach (1994a) studied ability judgments based on two dimensions (motivation and intelligence), each characterized by two scores, and showed that the configural intradimension relationships (the gap between the scores) were different than the configural interdimension relationships. These results suggest that the study of nonlinearity may be improved by incorporating information about attribute dimensionality.

The version of the scatter model that takes into account the impact of the dimension structure on configurality is labeled the *multiple-scatter* version. This version is expressed as

$$Y = \alpha + \sum_{i=1}^{m} \beta_i X_i + \sum_{j=1}^{n} \gamma_j \mathrm{SF}_j + \delta\, \mathrm{SFT}, \tag{2}$$

where m is the number of attributes, n is the number of dimensions, SF_j is the intradimension scatter of Dimension j, and SFT is the interdimension scatter.

SF_j is defined as

$$\mathrm{SF}_j = \left[\sum_{k=1}^{q_j} (X_k - F_j)^2 \right]^{1/2}, \tag{3}$$

where q_j is the number of attributes associated with Dimension j (the summation is on the attributes associated with Dimension j), and F_j is the mean of the q_j attributes associated with Dimension j.

SFT is defined as

$$\mathrm{SFT} = \left[\sum_{j=1}^{n} (F_j - \bar{F})^2 \right]^{1/2}, \tag{4}$$

where $\bar{F}$ is the mean of the dimensions, defined as

$$\bar{F} = \frac{1}{n} \sum_{j=1}^{n} F_j.$$

In these equations, the F_js are standardized, and if they have a negative correlation with the judgment, they are rescaled to have a positive correlation.

Analyses

A factor analysis of the three data sets showed that only the class-evaluation data set had a simple factor structure that was appropriate for an analysis by the multiple-scatter model. Therefore, in this section, this data set was analyzed by the multiple-scatter version.

A principal-components analysis with a varimax rotation of the class-evaluation data set revealed two dimensions. The first dimension was associated with attributes (the first five attributes) that were primarily related to the course (and were less related to the professor), and the second dimension was associated with attributes (the last five attributes) that were directly related to the professor. Indeed, in the class-evaluation questionnaire, the first five attributes are labeled *the level of the course,* whereas the last five attributes are labeled *the level of the teaching.*

To estimate the multiple-scatter version, the value of each of the two dimensions was determined by averaging across the scale scores associated with it; the value of each intradimension scatter was determined from the deviations of the dimension attributes from their mean (Equation 3); and the value of the interdimension scatter was determined from the deviations of the dimension values from their mean (Equation 4). Subsequently, both course evaluation and professor evaluation were modeled using Equation 2. That is, both were modeled as a function of the ten attributes, the interdimension scatter (SFT), the ten intradimension scatters, the scatter associated with the first dimension (SF_1), and the scatter associated with the second dimension (SF_2).

The scatter coefficients of the professor model and of the course model, as well as the standard errors of the estimates, are presented in columns 2 and 4 of Table 3. These data indicate that the coefficient of SF_2 was significantly negative ($p < .0001$) in both the professor model and the course model. However, whereas in the professor model the coefficient of SF_1 and the coefficient of SFT did not differ significantly from zero ($p > .1$), in the course model both coefficients were significantly negative ($p < .0005$ and $p < .0001$, respectively). Thus, whereas the simple version of the scatter model did not reveal any important differences in nonlinearity between the evaluation of professors and the evaluation of courses (the partial correlation between scatter and evaluation was quite similar in the two evaluation tasks; see column 2 of Table 1), substantial differences did appear when the data were analyzed by the multiple-scatter version.

What is the reason for these differences in nonlinearity? First, the difference between the two models with regard to the intradimension scatter coefficients is due to the fact that the inconsistency (i.e., scatter) associated with the first dimension is primarily relevant to the evaluation of the course, whereas the inconsistency

Table 3
Scatter Coefficients of Professor and Course Multiple-Scatter Models

Term	Professor model: Without course evaluation	Professor model: With course evaluation	Course model: Without professor evaluation	Course model: With professor evaluation
SF_1	−.007	.009	−.050	−.047
	(.012)	(.011)	(.013)	(.013)
SF_2	−.200	−.173	−.088	−.004
	(.140)	(.013)	(.017)	(.016)
SFT	−.022	.013	−.112	−.103
	(.140)	(.013)	(.016)	(.015)

Note. Numbers in parentheses are standard errors. SF_1 = intradimension scatter associated with the first dimension; SF_2 = intradimension scatter associated with the second dimension; SFT = interdimension scatter.

associated with the second dimension is primarily relevant to the evaluation of the professor. Second, the difference between the professor model and the course model with regard to the interdimension scatter coefficient is most likely due to the tendency to assign higher weights to positive (vs. negative) information when evaluating human objects than when evaluating nonhuman objects (Ganzach, 1993). This difference between the evaluation of human objects and nonhuman objects is associated with the *person positivity bias* (Sears, 1983), the tendency to evaluate human objects more favorably than nonhuman objects. Indeed, Sears's original finding that the overall evaluation of professors was more favorable than the overall evaluation of courses was replicated in my data, in which both evaluations were given on the same scale. The mean course evaluation in my data was 15.01 (SD = 2.61), whereas the mean professor evaluation was 15.49 (SD = 3.01). The difference between them was significant at $p < .0001$.

The differential impact of the various scatters on the evaluation of professors and on the evaluation of courses can be demonstrated by adding the course evaluation to the professor model and the professor evaluation to the course model. Columns 3 and 5 of Table 3 present the relevant scatter coefficients and their standard errors. It is clear from Table 3 that when the professor evaluation was added to the course model, the coefficient of SF_2 became nonsignificant ($p > .1$), whereas the coefficients of SF_1 and SFT changed very little. In contrast, it is also clear that adding the course evaluation to the professor model had little influence on the scatter coefficients. The path model that summarizes these results is given in Figure 1. This model suggests that interdimension inconsistency and inconsistency in the first dimension have an effect on course evaluation but have no effect on professor evaluation and that inconsistency in the second dimension has an effect on professor evaluation but has no (direct) effect on course evaluation. However, inconsistency in the second dimension has an indirect effect on course evaluation, which is mediated by the effect of professor evaluation on course evaluation.

In summary, the preceding analyses show that the simple version of the scatter model may be only a crude representation of nonlinearity in performance evaluation. Although the central feature of this nonlinearity, that is, conjunctive integration, is captured by this model, the intricate relationships between the inter- and intradimension scatters and the evaluated objects are captured by only the multiple-scatter model. This model shows that conjunctive integration depends on the particular dimension and the particular object.

Nonlinearity and Inconsistency Resolution

Inconsistency resolution (Ganzach, 1994a) refers to a decision strategy by which inconsistency between two or more aspects of the input information is resolved on the basis of other aspects of the information (see Lynch & Offir, 1989, and Slovic, 1966, for examples of inconsistency resolution). In particular, an (initial) overall evaluation may affect the weight of positive versus negative attributes (Ganzach, 1993). If this evaluation is high (low), relatively more weight is given to positive (negative) information than to negative (positive) information (Ganzach, 1994a). To demonstrate the implication of this type of inconsistency resolution for the relationship between scatter and rating, consider the following two pairs of job candidates judged on the basis of two equally important test scores. One pair consisted of two candidates with similar and high mean scores, which differed in the discrepancy between the two scores. The other pair consisted of two candidates with similar and low mean scores, which also differed in the discrepancy. Inconsistency resolution suggests that in the first pair, the candidate with the higher discrepancy would receive a higher rating. (If the resolution were based on the fact that

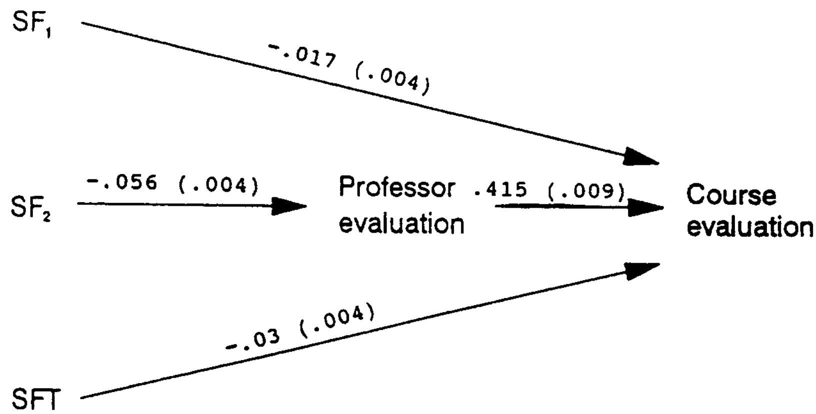

Figure 1. A schematic path diagram representing the effects of the inter- and intradimension scatters on course evaluation and on professor evaluation. The numbers above the arrows are the path coefficients, and the numbers in parentheses are their standard errors. Note that the diagram presents only the effects that are relevant to the analysis reported in the text (the nonlinear effects). SF_1 = intradimension scatter associated with the first dimension; SF_2 = intradimension scatter associated with the second dimension; SFT = interdimension scatter.

the mean score is high, the higher score would receive heavier weight.) This model also suggests that in the second pair, the candidate with the higher discrepancy would receive a lower rating. (If the resolution were based on the fact that the mean score is low, the lower evaluation would receive heavier weight.) Thus, inconsistency resolution suggests that the relationship between scatter and rating would be more positive for strong candidates than for weak candidates. This is demonstrated below with both the scatter of the simple version and the scatters of the multiple-scatter version.

Effect of Overall Evaluation on the Relationship Between the Scatter of the Simple Version and Rating

To investigate this effect, I divided each of the three data sets into two parts by a median split. One part (the "high" half) contained the targets that received high evaluations, and the other part (the "low" half) contained the targets that received low evaluations. (In the class-evaluation data set, two median splits were performed, one for professor evaluations and the other for course evaluations.) As expected from inconsistency resolution, the data indicate that the relationship between scatter and evaluation was more positive when the level of evaluation was high than when it was low. Columns 3 and 4 of Table 1 present the partial correlations between scatter and rating, controlling for attribute values. The inconsistency resolution hypothesis is strongly supported by these results. For three of the tasks, the partial correlation between scatter and evaluation was significantly negative in the low halves ($p < .0001$) and was nonsignificant in the high halves. In the professor-evaluation task, the effect was even more dramatic. Whereas the correlation between scatter and evaluation was significantly negative in the low half ($p < .0001$), it was significantly positive in the high half ($p < .0001$). Thus, although the relationship between scatter and evaluation may have been negative for the entire sample (see the analyses that were based on the scatter version), a more refined analysis revealed that this relationship was significantly negative only for weak targets. For strong targets, this relationship was not significantly negative, and it may even have been positive.

Effect of Overall Evaluation on the Relationships Between the Scatters of the Multiple-Scatter Version and Rating

Inconsistency resolution is also apparent when intra- and interdimension scatters are correlated with rating. The partial correlations (controlling for elevation) between the ratings and the three scatters are presented in Table 4. It is apparent that for the low halves all the partial correlations were strongly negative, whereas for the high halves most of these correlations tended to be positive. Again, the relationships between the various scatters and overall evaluation suggested by these analyses were more complex than the relationships suggested by the previous analyses based on the multiple-scatter version.

The strength of inconsistency resolution can be estimated from the difference between the partial correlation in the high half and the partial correlation in the low half (the partial correlation difference). First, note that inconsistency resolution, as described above, suggests that

Table 4
Partial Correlations Between Rating and the Three Scatter Terms of the Multiple-Scatter Model

	Entire data set			High half			Low half		
Task	SF_1	SF_2	SFT	SF_1	SF_2	SFT	SF_1	SF_2	SFT
Professor evaluation	−.042**	−.123**	−.035**	−.067**	.083**	.106**	−.144**	−.231**	−.173**
Course evaluation	−.061**	−.064**	−.076**	.031*	−.027*	.013	−.156**	−.089**	−.139**

Note. SF_1 = intradimension scatter associated with the first dimension; SF_2 = intradimension scatter associated with the second dimension; SFT = interdimension scatter.
* $p < .01$. ** $p < .0001$.

the partial correlation difference would be positive. The data indicate that this was indeed the case. For SF_1, the differences were .077 and .187 for professor evaluation and for course evaluation, respectively. For SF_2, the differences were .314 and .062, respectively. For SFT, the differences were .279 and .152, respectively.

Second, because the first dimension is more relevant to course evaluation and the second dimension is more relevant to professor evaluation, it is likely that with regard to SF_1, inconsistency resolution would be more pronounced in the evaluation of courses, but with regard to SF_2, inconsistency resolution would be more pronounced in the evaluation of professors. Because the partial correlation difference represents the strength of inconsistency resolution, this prediction can be tested by examining the magnitude of the partial correlation difference for each of the two scatters and each of the two evaluation tasks. Indeed, the data indicate that the partial correlation difference was higher for course evaluation than for professor evaluation with regard to SF_1 and was higher for professor evaluation than for course evaluation with regard to SF_2.

Discussion

The results of the analyses of the three data sets indicate that in a variety of real-world performance evaluation tasks, rating does not follow a linear rule. Substantial nonlinear effects did exist in all three data sets that were examined.

In agreement with previous laboratory work, there is a basic tendency for performance evaluations to be conjunctive, that is, to exhibit a negativity bias. However, it is important to note that this basic tendency may be mitigated by a number of factors, such as whether the evaluated object is human or nonhuman (e.g., professor or course) and whether the general evaluation of the object is positive or negative (e.g., good or bad professor; see Ganzach, 1993, for other examples). These conditions could be further explored in future research. For example, on the basis of research on the effect of justification of rating on the level of rating (Decker & Cornelius, 1981), it could be hypothesized that the tendency to be conjunctive or disjunctive may be influenced by a requirement to justify the rating.

A number of nonlinear models were used in previous research to detect nonlinear strategies in rating, and disjunctive–conjunctive strategies in particular. Einhorn's (1970) hyperbolic and parabolic models are the most well-known nonlinear models, but other models were also examined (e.g., Brannick and Brannick, 1989; Goldberg, 1969, 1971; Johnson et al., 1989; Weldon & Mustari, 1988; Wiggins & Hoffman, 1968). Although these models did provide some evidence for reliance on nonlinear rules, they failed to detect systematic nonlinear rules in judgments that were based on nonorthogonal stimuli (Goldberg, 1971). The scatter model fares better than these alternative models because it is more powerful in detecting nonlinearity if the underlying strategies are indeed conjunctive or disjunctive (Ganzach & Czaczkes, 1995).

Finally, what are the practical implications of this research? It is clear that the analysis of rating with the scatter model can supply raters with insights that they may not have (Nisbett & Wilson, 1977) about their tendency to be disjunctive or conjunctive. However, are these insights important in view of the meager incremental variance explained by the scatter terms? My view is that this is still an important insight, because raters often face the task of rating a set of alternatives that are roughly equal in value. These sets are characterized by a lack of dominant alternatives and negative interattribute correlations. In such sets, the difference between linear and nonlinear models may be quite large (Curry, Louviere, & Augustine, 1981).

The understanding of the nonlinear elements of rating is particularly important when some information about the relationship between the attributes and the criterion does exist, and raters could be provided with information about the gap between the nonlinear aspects of their rating and nonlinearity in the environment. Although some evidence does exist that nonlinearity characterizes judgment more than it characterizes the environment (Brehmer, 1972), the question of whether the strong nonlinearity that exists in performance evaluation exists in the environment as well is open for future research.

References

Bernadin, H. J., & Vilanova, P. (1986). Performance appraisal. In E. A. Lock (Ed.), *Generalizing from laboratory to field setting* (pp. 43-62). Lexington, MA: Lexington Books.

Blom, G. (1958). *Statistical estimates and transformed beta variables.* New York: Wiley.

Bolster, B. I., & Springbett, B. M. (1961). The reaction of interviewers to favorable and unfavorable information. *Journal of Applied Psychology, 45,* 97–103.

Brannick, M. T., & Brannick, J. P. (1989). Nonlinear and noncompensatory processes in performance evaluation. *Organizational Behavior and Human Decision Processes, 44,* 97–122.

Brehmer, B. (1972). Cue utilization and cue consistency in multiple-cue probability learning. *Organizational Behavior and Human Decision Processes, 8,* 286–296.

Campbell, J. P. (1986). Labs, fields and straw issues. In E. A. Lock (Ed.), *Generalizing from laboratory to field setting* (pp. 269–279). Lexington, MA: Lexington Books.

Cronbach, L. J., & Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin, 68,* 104–109.

Curry, D. J., Louviere, J. J., & Augustine, M. (1981). Comment on the sensitivity of brand choice to attribute importance weights. *Decision Sciences, 12,* 502–516.

Dawes, R. M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin, 81,* 95–106.

Decker, P. J., & Cornelius, E. T. (1981). The effect on leniency of justifying performance ratings to a supervisor. *Journal of Psychology, 108,* 211–218.

DeNisi, A. S., Cafferty, T. P., & Meglino, B. M. (1984). A cognitive view of the performance appraisal process: A model and research propositions. *Organizational Behavior and Human Decision Processes, 33,* 360–396.

Einhorn, H. J. (1970). The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin, 73,* 221–230.

Einhorn, H. J. (1971). The use of nonlinear, noncompensatory models as a function of task and amount of information. *Organizational Behavior and Human Decision Processes, 6,* 1–27.

Einhorn, H. J., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Linear regression and process tracing models of judgment. *Psychological Review, 86,* 464–485.

Fusilier, M. R., & Hitt, M. A. (1983). Effects of age, sex, race, and employment experience on students' perceptions of job candidates. *Perceptual and Motor Skills, 57,* 1127–1134.

Ganzach, Y. (1993). Goals as determinants of nonlinear, noncompensatory judgment strategies. *Organizational Behavior and Human Decision Processes, 56,* 422–440.

Ganzach, Y. (1994a). Inconsistency and uncertainty in multiattribute judgment of human performance. *Journal of Behavioral Decision Making, 7,* 193–211.

Ganzach, Y. (1994b). Theory and configurality in expert and layperson judgment. *Journal of Applied Psychology, 79,* 439–448.

Ganzach, Y., & Czaczkes, B. (1995). On detecting nonlinear, noncompensatory strategies in judgment: Comparison of alternative regression models. *Organizational Behavior and Human Decision Processes, 61,* 168–177.

Goldberg, L. R. (1969). The search for configural relationships in personality assessment: The diagnosis of psychosis vs. neurosis from the MMPI. *Multivariate Behavioral Research, 4,* 523–536.

Goldberg, L. R. (1971). Five models of clinical judgment: An empirical comparison between linear and nonlinear representation of the human inference process. *Organizational Behavior and Human Performance, 6,* 458–479.

Goldberg, L. R. (1976). Man versus model of man: Just how conflicting is the evidence? *Organizational Behavior and Human Performance, 16,* 13–22.

Guion, R. M. (1991). Personnel assessment, testing, and placement. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (pp. 327–397). Palo Alto, CA: Consulting Psychologists Press.

Guion, R. M., & Gibson, W. M. (1988). Personnel selection and placement. *Annual Review of Psychology, 39,* 349–374.

Hamilton, D. L., & Huffman, L. J. (1971). Generality of impression-formation processes for evaluative and nonevaluative judgments. *Journal of Personality and Social Psychology, 20,* 200–207.

Hitt, M. A., & Barr, S. H. (1989). Managerial selection decision models: Examination of configural cue processing. *Journal of Applied Psychology, 74,* 53–61.

Hollman, T. D. (1972). Employment interviewers' errors in processing positive and negative information. *Journal of Applied Psychology, 56,* 130–134.

Ilgen, D. R., & Favero, J. L. (1985). Limits in generalization from psychological research to performance appraisal processes. *Academy of Management Review, 34,* 249–262.

Jago, A. G. (1978). Configural cue utilization in implicit models of leader behavior. *Organizational Behavior and Human Decision Processes, 22,* 474–496.

Johnson, E., Meyer, R. J., & Ghose, S. (1989). When choice models fail: Compensatory choice models in negatively correlated environment. *Journal of Marketing Research, 6,* 255–270.

Kanouse, D. E., & Hanson, L. R. (1972). *Negativity in evaluation.* New York: General Learning Press.

London, M., & Hakel, M. D. (1974). Effects of applicant stereotypes, order, and information on interview impressions. *Journal of Applied Psychology, 59,* 157–162.

London, M., & Poplawski, J. R. (1976). Effects of information on stereotype development in performance appraisal and interview contexts. *Journal of Applied Psychology, 61,* 199–205.

Lynch, J. G., & Offir, C. (1989). Effects of cue consistency and value on base rate utilization. *Journal of Personality and Social Psychology, 56,* 170–181.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Ogilvie, J. R., & Schmitt, N. (1979). Situational influences of linear and nonlinear use of information. *Organizational Behavior and Human Performance, 23,* 292–306.

Sears, D. O. (1983). The person positivity bias. *Journal of Personality and Social Psychology, 44,* 233–250.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin, 105,* 131–142.

Slovic, P. (1966). Cue consistency and cue utilization in judgment. *American Journal of Psychology, 79,* 427–434.

Stumpf, S. A., & London, M. (1981). Capturing rater policies in evaluating candidates for promotion. *Academy of Management Journal, 24,* 752–766.

Tukey, J. W. (1962). The future of data analysis. *Annals of Mathematical Statistics, 33,* 22.

Webster, E. C. (1964). *Decision making in the employment interview.* Montreal, Quebec, Canada: Eagle.

Weldon, E., & Mustari, E. L. (1988). Felt dispensability in groups of coactors: The effects of shared responsibility and explicit anonymity on cognitive effort. *Organizational Behavior and Human Decision Processes, 41,* 330–351.

Wiggins, N., & Hoffman, P. J. (1968). Three models of clinical judgment. *Journal of Abnormal Psychology, 73,* 70–77.

Wyer, R. S., & Hinkle, R. L. (1976). Informational factors underlying inferences about hypothetical people. *Journal of Personality and Social Psychology, 34,* 481–495.

Received September 6, 1994
Revision received January 5, 1995
Accepted January 5, 1995 ■