

Performance Evaluation in a Dynamic Context: A Laboratory Study of the Impact of a Prior Commitment to the Ratee

Max H. Bazerman

School of Management, Boston University

Rafik I. Beekun

Graduate School of Industrial Administration
Carnegie-Mellon University

F. David Schoorman

University of Maryland

A dynamic view of performance evaluation is proposed that argues that raters who are provided with negative performance data on a previously promoted employee will subsequently evaluate the employee more positively if they, rather than their predecessors, made the earlier promotion decision. A total of 298 business majors participated in the study. The experimental group made a promotion decision by choosing among three candidates, whereas the control group was told that the decision had been made by someone else. Both groups evaluated the promoted employee's performance after reviewing 2 years of data. The hypothesized escalation of commitment effect was observed in that the experimental group consistently evaluated the employee more favorably, provided larger rewards, and made more optimistic projections of future performance than did the control group.

After a long history of attempting to identify optimal performance appraisal instruments and techniques (cf. Landy & Farr, 1980), recent research has moved towards studies that describe the processes involved in making a judgment about performance (Borman, 1975; Feldman, 1981; Schneider, 1977; Zedeck & Kafry, 1977). This article represents this latter paradigm and specifically addresses the dynamic processes that occur within a rater across multiple judgments of the same ratee. It is concerned with a longitudinal view of the rating process in order to focus on judgmental demands associated with rating the same ratee at multiple points in time. This article presents: (a) a brief review of the trend towards understanding rating processes, (b) presentation of a longitudinal viewpoint, (c) development of a set of hypotheses from this new viewpoint, and (d) discussion of the results of the empirical test of these hypotheses.

The traditional view of performance evaluation assumes that a rater has both a complete set of information about the behavior of the ratee and an organizationally provided performance appraisal instrument. The instrument provides the

rater with the questions to be answered about a particular ratee. The rater searches through the information available, identifies all behaviors pertinent to the questions to be answered, integrates all this information in a summary evaluation, and provides the rating response that most closely approximates the summary evaluation. From a psychometric viewpoint, such evaluations can be represented as consisting of a true score and an error term (Nunnally, 1967). The error term can be dichotomized into systematic and nonsystematic components. The latter have typically been considered to be randomly distributed, whereas the systematic portion has been seen to result from certain biases of the rater (e.g., the halo effect).

Recent research has sought to move beyond the identification of these systematic biases by attempting to provide process models that explain why these biases occur. Attention has focused on identifying the information acquisition, information integration, and judgmental process that raters perform in making an evaluation. A primary goal of this research direction has been to determine the processes within the rater that lead to error in the evaluation system. Such identification may be potentially useful in training evaluators to eliminate error.

Although recent attempts at developing a cognitive model of the rater have greatly improved our understanding of the performance evaluation process, these attempts have been made within a static perspective. Both the traditional view of performance evaluation and more recent cogni-

The authors thank Lloyd Baird and two anonymous reviewers for critical comments on earlier drafts of this paper.

Requests for reprints should be sent to Max H. Bazerman, 621 Commonwealth Avenue, School of Management, Boston University, Boston, Massachusetts 02215.

tive models have examined how a rater evaluates an employee at one point in time. Realistically, however, a common rating pattern consists of an individual making repeated ratings of a ratee at multiple points in time. Psychologically, this creates a new set of demands on the rater—to rate individuals objectively at a given point in time without being affected by his or her previous judgments of the same ratee. A bit of reflection suggests that obtaining such objectivity is a formidable task at best and may be unattainable.

From a decision-making perspective, a number of researchers (Staw, 1976, 1981; Bazerman, Schoorman, & Goodman, Note 1; Conlon & Wolf, Note 2) have demonstrated that decision makers who commit themselves to a particular course of action may make subsequent, related decisions in a nonoptimal manner in order to justify the previous commitment. For example, Staw (1976) examined the tendency to escalate commitment of financial resources following negative feedback concerning a previous related allocation. One group of subjects was requested to allocate research and development funds to one of two operating divisions of an organization. These subjects were then told that, over the 3-year period following their initial decision, their investment decision proved unsuccessful and that they were faced with a second allocation decision concerning the division to which they had given funds previously. A control group was told of a decision made by another financial officer of the firm that had not been successful (the same content information as provided to the previous group of subjects) and that they were to make the second allocation of funds concerning that division. It was found that subjects allocated significantly more funds to the original division in the second allocation decision when they, rather than another financial officer of the firm, had made the initial allocation decision.

These findings may provide some insight into rater biases in the performance appraisal context. Consider, for example, the following scenario: A rater evaluates multiple individuals for a promotion. After giving serious consideration to a number of candidates, the rater selects one individual. After 2 years in the new role, the promoted individual has not lived up to expectations and there exists little evidence that his or her performance will dramatically improve in the near future. The "rational" rater will evaluate the employee poorly and may consider a personnel change. However, if the rater had made the initial promotion decision, the subsequent decision may be biased for at least three reasons. First, the rater's perception of information may be biased by his or her previous decision. That is, the rater may

pay greater attention to information that supports his or her promotion decision than to information that discredits the earlier decision. Second, given the set of information that the rater perceives, he or she may make biased judgments due to the prior decision. For example, the evaluator may give greater weight to confirmatory than disconfirmatory information (Nisbett & Ross, 1980; Snyder & Swann, 1978; Lepper, Ross, & Lau, Note 3). Third, even if the rater privately judges the ratee's performance in a negative manner, he or she may not give a negative evaluation since such action may publicly contradict the previous commitment (i.e., the promotion) to the ratee. Overall it is predicted that raters who make an initial commitment (e.g., a promotion) to an employee will systematically make subsequent related evaluations of that employee in a biased manner in order to justify the initial support of the ratee.

The current investigation examines the generalizability of the escalation hypothesis to the performance appraisal situation. In past escalation research, subjects have made initial financial commitments to an organization's decision(s). Research suggests that subjects will demonstrate the tendency to escalate when asked to make a very similar decision at multiple points in time. Performance appraisal, however, provides a new set of questions, since the rater is making judgments about promotability or hiring at one time and judgments about performance at a later time. Also, if the performance evaluation judgment is affected by the prior promotion decision, will this bias generalize to other evaluation decisions such as recommendations for salary increases and/or predictions of future performance? It is anticipated that the escalation effect will be observed in the performance appraisal judgment, and that the bias will generalize to related decisions. Based on the preceding discussion, the central proposition of the study may be stated as follows: A rater is more likely to evaluate (reward, assess promotion potential, forecast future performance) a poorly performing ratee higher when he or she previously promoted the ratee to his or her current position than when another rater made the initial promotion decision.

Method

Two hundred and ninety-eight business majors at the University of Texas at Austin participated in a decision making exercise as part of a course. Subjects were randomly assigned to one of four experimental conditions in a 2×2 design. Each subject was presented with the "Heeley's Stores Case" in which he or she was asked to play the

role of a vice-president who was either: (a) promoting and later evaluating an executive (high-responsibility condition); or (b) only evaluating an executive who had already been promoted by someone else (low-responsibility condition).

Subjects were told that the purpose of the case was to study performance appraisal procedures. In order to maximize the subjects' involvement, the experimenter told them that their decisions would help the federal district court of a nearby city in its investigation of the procedures followed by private businesses in promotion evaluations.

The Heeley's Performance Appraisal Case

High-responsibility condition. The performance appraisal case presented to high-responsibility (HR) subjects describes a hypothetical corporation (Heeley's) in 1975. A scenario is presented in which the vice-president of a corporation, whose role the subject is to assume, must appraise the performance of three merchandise managers.

The case provides three hypothetical sets of materials concerning each manager—the three are the best Region 5 managers for Heeley's. A brief description of the managers is followed by data (including diagrams) that includes sales, earnings, return on investment, inventory turnover ratios, and part performance ratings. Those subjects were then asked to decide which of the three employees should be promoted to be regional director in the organization. Besides writing down the name of the new regional director, the HR subjects were also asked to write a brief paragraph explaining why they selected that manager. The HR subjects were then administered a second section of the case entitled *Two Years Later*. The new material provided negative data with regard to the new regional director's performance during his first 2 years in office. The subject was then asked to make a number of evaluations (the dependent variables).

Low-responsibility condition. The low-responsibility (LR) subjects received the same set of materials concerning the three managers, but were informed that the former vice-president (their predecessor) had appointed one of the three executives as regional director. An equal number of LR subjects were randomly assigned one of the three managers as the promoted regional director and were provided with subsequent negative data regardless of who had been appointed. The LR subjects were then asked to make the same evaluations as those made by HR subjects.

Dependent Variable

The dependent variable used was the subjects' commitment to a previously promoted employee as demonstrated by multiple performance appraisal assessments. Three groups of dependent measures were obtained: rewards, appropriateness of promotion/demotion, and forecasted future performance. Rewards were measured by asking subjects to (a) set the percentage pay increase for the former employee (company average = 10%), and (b) designate the number of bonus vacation days to give to this employee (minimum = 0, maximum = 4). Promotion/demotion potential was measured by three 11-point Likert items that assessed the former employee's promotability within the next 5 years, current appropriateness for a demotion, and current appropriateness for a layoff if 15% of the company's managerial positions were eliminated. Finally, forecasted future performance was measured by asking subjects to estimate sales and returns for the former employee's region for the upcoming 3-month period. All seven measures were coded such that a higher number indicated a higher evaluation.

Results

A preliminary analysis examined whether the choice among the three potential managers had

Table 1
Means and Standard Deviations of Dependent Variables for Low- and High-Responsibility Subjects

Dependent variable	Low responsibility (<i>n</i> = 160)		High responsibility (<i>n</i> = 138)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Pay increase	8.87	2.88	9.72	3.37
Vacation increase	1.39	.99	1.83	1.15
Promotion potential	3.87	2.09	4.56	2.35
Demotion potential	5.89	2.38	6.99	2.31
Layoff decision	4.79	2.55	5.51	2.58
Return projection	\$ 702,180	\$ 92,858	\$ 726,937	\$ 70,304
Sales projection	\$2,900,463	\$270,882	\$2,944,542	\$219,733

an impact on the evaluations made at the second decision point. If there were main effects for this choice, it would be necessary to include this choice as an independent variable. The choice of employee did not significantly affect any of the seven dependent measures. In addition, there was no clear pattern in the direction of nonsignificant effects across the measures.

The effects of the experimental manipulation of responsibility on the seven dependent variables were analyzed using a multivariate analysis of variance procedure. The Wilks's lambda criterion was significant for responsibility, $F(7, 288) = 3.85$, $p < .001$.

Table 1 shows the cell means for each of the seven variables. HR subjects gave higher pay increases, $F(1, 294) = 5.33$, $p < .05$; more vacation days, $F(1, 294) = 13.320$, $p < .001$; and more positive evaluations of promotion potential $F(1, 294) = 7.269$, $p < .001$. They were less likely to demote the former employee, $F(1, 294) = 16.149$, $p < .001$; or layoff the employee, $F(1, 294) = 5.738$, $p < .05$; and forecasted higher future return, $F(1, 294) = 6.522$, $p < .05$. The forecast for sales, although in the predicted direction, was not statistically significant.

Discussion

The main effect of responsibility upon performance evaluations of a previously promoted employee could be explained by at least three processes. First, after promoting an employee, a rater's perceptual processes may be biased such that negative cues about the focal employee's performance may not be fully acquired. Second, even if the negative information is cognitively acquired, the subject's judgmental processes may be biased such that the new information is immediately discounted or later given little weight in making the final evaluation. Finally, even if the negative information is acquired and a nonbiased judgment is made, the subject may not be willing to publicly contradict his or her prior decision. Although all three explanations are viable and can explain the effect, we need further research to delineate which of these actually accounts for the escalation finding.

The current findings concerning the impact of prior commitment to a previously promoted employee have a number of important implications. This research introduces a new bias to the performance evaluation literature. It also points out the need for a dynamic perspective of the performance appraisal process. In addition, the current results add support to the escalation paradigm by demonstrating the effect in a new, nonfinancial context. Finally, our findings provide evidence of the generalizability of a decision maker's bias due to a prior commitment. That is, the results indi-

cate that an initial promotion decision subsequently affects reward and forecasting decisions, as well as evaluations of promotability.

Future research on performance evaluation should continue to examine rater cognitions as a source of bias in the evaluation process. More specifically, added research is needed that identifies rater cognitions in their typical dynamic context. Finally, longitudinal field research is needed to increase the generalizability of findings concerning the dynamic role of the cognitions of raters.

Reference Notes

1. Bazerman, M. H., Schoorman, D. F., & Goodman, P. S. *A cognitive evaluation of escalation processes in managerial decision-making*. Paper presented at the 40th annual meeting of the Academy of Management, Detroit, August 1980.
2. Conlon, E. J., & Wolf, G. *Escalation as sunk costs: An environmental and cognitive analysis*. Paper presented to the 40th annual meeting of the Academy of Management, Detroit, August 1980.
3. Lepper, M. R., Ross, L., & Lau, R. *Persistence on inaccurate and discredited personal impressions: A field demonstration of attributional perseverance*. Unpublished manuscript, Stanford University, 1979.

References

- Borman, W. C., & Dunnette, M. D. Behavior-based vs. trait-oriented performance ratings: An empirical study. *Journal of Applied Psychology*, 1975, 60, 561-565.
- Feldman, J. M. Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology*, 1981, 66, 127-148.
- Landy, F. J., & Farr, J. L. Performance rating. *Psychological Bulletin*, 1980, 87, 72-107.
- Nisbett, R. E., & Ross, L. *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, N.J.: Prentice-Hall, 1980.
- Nunnally, J. C. *Psychometric Theory*. New York: McGraw-Hill, 1967.
- Schneider, C. Operational utility and psychometric characteristics of behavioral expectation scales: A cognitive reinterpretation. *Journal of Applied Psychology*, 1977, 62, 541-548.
- Snyder, M., & Swann, W. B. Behavioral confirmation in social interaction: From social perception to social reality. *Journal of Experimental Social Psychology*, 1978, 14, 148-162.
- Staw, B. M. Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance*, 1976, 16, 27-44.
- Staw, B. M. The escalation of commitment to a course of action. *Academy of Management Review*, 1981, 6, 577-587.
- Zedeck, S., & Kafry, D. Capturing rater policies for processing evaluation data. *Organizational Behavior and Human Performance*, 1977, 18, 269-294.

Received April 12, 1982

Revision received April 30, 1982 ■