

UC Irvine

UC Irvine Previously Published Works

Title

Employee Responses to Formal Performance Appraisal Feedback

Permalink

<https://escholarship.org/uc/item/8hj58186>

Journal

Journal of Applied Psychology, 71(2)

ISSN

0021-9010

Authors

Pearce, Jone L
Porter, Lyman W

Publication Date

1986-05-01

DOI

10.1037/0021-9010.71.2.211

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Employee Responses to Formal Performance Appraisal Feedback

Jone L. Pearce and Lyman W. Porter
Graduate School of Management, University of California, Irvine

The present study investigates the attitudinal impacts of the receipt of formal performance appraisal feedback. It is suggested that the feedback that one is "satisfactory" will be disconfirming for many feedback recipients. Therefore, it is hypothesized that (a) attitudes toward the performance appraisal systems and (b) organizational commitment will decrease and remain lower for those receiving "satisfactory" ratings, whereas the attitudes of those receiving higher appraisal ratings will remain unchanged. The hypotheses are tested on panels of management and nonmanagement employees (these latter receiving new appraisals 12 months after their managers) in two federal agencies over a 30-month period using perceived and actual performance ratings. There was a significant and stable drop in the organizational commitment of satisfactory employees after the introduction of formal appraisals, with mixed results for attitudes toward the appraisal system. The findings suggest that potentially negative consequences of implicitly comparative formal performance appraisals can occur for those performing at a satisfactory, but not outstanding, level. This study also provides an empirical check on the accuracy of self-reported appraisal ratings.

Performance appraisal is one of the most widely researched topics in all of personnel psychology. In recent years attention has been even greater because of important potential implications relating to fair employment practices and because of increasing concerns about employee productivity in organizations. Much of this rather voluminous literature on appraisals has focused on improving the accuracy of ratings by means of better instrumentation and more effective rater training. However, the present study is concerned with the effects of received appraisals on those being evaluated.

Formal performance appraisals can be viewed as a particular kind of feedback. One of the primary purposes of formal appraisals is the provision of clear, performance-based feedback to employees (Carroll & Schneier, 1982). There has been a substantial body of research on the effects of feedback, with Ilgen, Fisher, and Taylor (1979) providing a comprehensive review of the research literature. They identify the *sign* of the feedback—whether it is seen as positive or negative—as one of the key variables in message perception. In fact, Landy and Farr (1983) view the sign of the message as "the most important message characteristic in terms of its impact on the acceptance of feedback" (p. 168).

Despite the availability of prior research concerning the "acceptability" of positive/negative feedback, no empirical studies were located that have directly investigated the impact of the sign of the feedback on the subsequent attitudes and behavior of employees in the work situation. Although rigorously developed research on perceptions of feedback suggests that the sign of the feedback message is critical in perception of message con-

tent, it provides no data concerning the effects of feedback perceptions on subsequent attitudes and behaviors. There have been suggestions that performance appraisal feedback can have negative impacts on recipients' attitudes and subsequent behaviors, but these observations have not been subjected to empirical test.

Several theorists have, however, described "defensive" responses to negative feedback. De Nisi, Randolph, and Blencoe (1980) suggest that employees may attempt to retaliate when they receive low ratings from peers. Taylor, Fisher, and Ilgen (1984) argue that, even when a feedback system is perceived as fair, negative feedback may threaten employees' perceived freedom of choice and could result in defiant opposition or reaction to the supervisor.

Thompson and Dalton (1970) and Meyer (1975) have written on the dysfunctional effects of appraisal ratings feedback. Thompson and Dalton (1970) analyzed the experiences of several engineering firms with the introduction of formal appraisals for their technical employees. They found widespread dissatisfaction and reported generally lowered individual self-confidence and job performance. The authors noted:

Performance appraisal touches on one of the most emotionally charged activities in business life—the assessment of a man's contribution and ability. The signals he receives about this assessment have a strong impact on his self-esteem and on his subsequent performance. (Thompson & Dalton, 1970, p. 150.)

Similarly, Meyer (1975) suggested that employees given "below average" merit ratings become alienated and demoralized. This is because most employees consider their own work performance to be "above average." For example, Meyer (1975) reported the results of four samples of employees rating their own performance. He found that from 70%–80% of the employees in these samples rated their own performance as "in the top 25%." Meyer, Kay, and French (1965) found that criticism had a negative effect on the achievement of goals, whereas praise had little effect one way or the other. There is, however, a virtual absence of research

We would like to thank James L. Perry, Debora E. Dunkle, and Carla Hanzlik for their assistance.

Correspondence concerning this article should be sent to Jone L. Pearce, Graduate School of Management, University of California, Irvine, California 92717.

evidence to support these or any other conjectures concerning the possible reactions of employees to their appraisal ratings, and therefore the present study is concerned with testing these arguments.

This study was part of a large-scale research effort to evaluate the effects of the 1978 Civil Service Reform Act (Perry & Porter, 1981). One of the primary areas of the Act involved the initiation of formal objectives-based performance appraisal procedures in all Federal agencies. Although appraisal systems had previously been put into operation in some agencies prior to the Act, they were not conducted systematically and universally throughout the government. The Act mandated that formal performance-based procedures be developed and implemented in all agencies. In the larger investigation of which this study was one component, five diverse federal agencies formed the sample of organizations. For two of these five agencies it was possible to obtain a sufficient sample of employee attitude data before and after the imposition of the new performance appraisal systems. In each agency attitude data were collected from separate samples of managers and non-management employees, because each group received the new appraisal ratings at different times—managers 12 months before nonmanagement employees. This allowed limited control of the "history effects," with the lower level employees acting as a control group during the first 12 months of the managers' appraisal period.

Although actual performance appraisal ratings are available in only one of the agencies, self-reported ratings are available in both. However, because the present study is concerned with reactions to appraisal feedback, *perceived* rating would be expected to be a better predictor of recipient reactions and to represent the primary focus of this investigation. Nevertheless, both actual (where available) and self-reported ratings results are reported. Thus, one by-product of the present study is the opportunity to examine the accuracy of self-reported ratings in the one agency in which both are available.

The present sample—two federal agencies, each composed of managers receiving feedback on the new performance-based appraisal system 12 months before nonmanagement employees in the same agency—provides a unique opportunity to test the effects of performance appraisal feedback. It is rare in field settings to find the introduction of a new form of appraisal feedback for some groups but with other groups not receiving such feedback. For the first time in these organizations individuals were given a formal performance rating, based on objectives for their specific jobs, that was (by law) to be used for all promotions, pay increases, and reductions-in-force (layoffs). Furthermore, the data were available for all respondents over a 30-month period, before and after appraisal feedback, to both management and nonmanagement employees. Finally, the two agencies serve as independent tests (each can be seen as a replication of the test in the other) of the hypotheses.

It is suggested that performance feedback that one is "satisfactory" or "meeting standards" will be experienced as negative by many of these appraisal feedback recipients, not just by those receiving objectively poor performance ratings. This follows from Meyer (1975), among others, who indicated that most employees consider themselves to be above-average performers. Additional support is provided by Parker, Taylor, Barrett, and Martens (1959), Ilgen, Peterson, Martin, and Boesch (1981), and Smir-

cich and Chesser (1981), who found that subordinates rate their own performance more highly than their supervisors do. This discrepancy in the perceptions of supervisors and subordinates has also been noted by Mowday (1983) and Feldman (1981). Thus, it might be expected that many average performers would experience such feedback as negative.

If subordinates experience disconfirmation of their self-perceptions of "above average" performance, it could be expected to result in experienced cognitive dissonance. Such individuals should, therefore, be motivated to reduce the dissonance (Festinger, 1961). However, there is little evidence from employment situations to suggest how they might attempt to reduce it. Limited direction is provided by Mowday (1983), who suggests that ego-defensive bias would lead employees receiving perceived poor ratings to blame their failure on external factors rather than on their own personal characteristics.

One way to reduce dissonance is to minimize the importance of the information received. However, this is particularly difficult with appraisals that represent the organization's formal assessment of one's performance. As Thompson and Dalton (1970) noted above, appraisals become emotionally important to many in organizations. Furthermore, when appraisals will be used to administer organizational rewards such as promotions and pay (as in the present agencies), it is even more difficult to devalue them. Alternatively, blame can be placed on such external factors as task difficulty, the source of the rating, that is, the supervisor, or the organization that sponsored such a (poorly run) appraisal system. This follows from Meyer (1975), who suggested that employees given below-average merit raises often become disenchanted with their employers, and Thompson and Dalton (1970), who argued that any system that includes either explicit or implicit peer-comparisons results in self-blame, (lowered employee self-confidence and reduced individual performance) and employer blame (increased turnover).

The present hypotheses follow from this attribution theory framework. Following Mowday (1983) and others, we expect the ego defense mechanisms for such important feedback as work performance will lead those receiving (merely) satisfactory ratings to blame external factors rather than themselves. We would expect the most likely targets of such blame to be the supervisor (who made the judgment), the performance appraisal system (inappropriate measures), and the organization (it developed the policy), with no *a priori* predictions concerning which target will be preferred. There is some evidence that characteristics of the appraisal system (and other features of the feedback process) may influence which target is chosen. Landy, Barnes-Farrell, and Cleveland (1980) found that characteristics of the performance appraisal system (e.g., frequency of evaluation, supervisor's knowledge of performance) were better predictors of the perceived "fairness" of the evaluation than was the actual rating received by the employee. That is, when the appraisal system was perceived as fair and accurate, it was not blamed for the low rating. In the present study, those receiving relatively high ratings would be expected to experience "confirmation" of their positive self-perception, and thus we would expect no change in their attitudes toward these objects (Meyer et al., 1965). Data are available on two of these external targets. (Because the present test is taken from a larger evaluation of the impact of personnel changes on these organizations rather than a direct assessment of employee

attributional processes, no reliable scale for "blaming the supervisor" was used during the entire 30-month study period.) Therefore, the present study addresses two hypotheses:

Hypothesis 1. Attitudes toward the performance appraisal system will decrease and remain lower after the introduction of the appraisal system for those receiving relatively low ratings on it, whereas the attitudes of those receiving relatively high ratings will be unchanged.

Hypothesis 2. Those receiving a relatively low rating on the appraisal system will have more negative attitudes toward the organization after its introduction and the attitudes will remain significantly lower, whereas the attitudes of those with relatively high ratings will remain unchanged.

Method

Context and Sample

Data were obtained from a random sample of federal civil service employees from a National Aeronautic and Space Administration research center and a Department of Defense engineering station. The research center is responsible for research in the storing of lightweight flight structures, fluid mechanics, and fundamental aerodynamics. The engineering station provides inservice engineering support for the surface warfare systems of the United States Fleet and the ships of friendly nations. Engineers and scientists are the dominant employee groups in both organizations (engineers, scientists, and professional administrators comprise 64% of the respondents; technicians, 23%; and clerical and trades, 13%). Seventy-eight percent of the sample is male and 78% are non-Hispanic whites. The present study reports data from these two organizations for managers—managers and supervisors in federal civil service Grades 13 to 15—and nonmanagement employees (hereafter called "employees")—nonsupervisors in Grades 13 to 15 and all employees in Grades 12 and below. The data are analyzed separately, because the managers received their first ratings on the new performance appraisal system a year before the employees. Several other research sites that were part of the larger evaluation of the effects of the Reform Act could not be used in the present study because the number of managers in these agencies was prohibitively small for the analytic techniques employed.

Prior to the implementation of the Act's mandated objectives-based performance appraisal system, both agencies used more informal and subjective means to appraise performance. Appraisal practices at the research center varied widely. Many supervisors preferred informal face-to-face sessions with subordinates. One center respondent referred to them as "blank paper reviews," since the subordinate and supervisor would sit down to discuss performance, "each with a blank piece of paper." At the engineering station there was a more formal system, but it was "trait-based," in which supervisors rated their subordinates on characteristics such as "leadership."

Title II of the Civil Service Reform Act of 1978 mandated at least yearly evaluation on an appraisal system: "establishing performance standards which will, to the maximum extent feasible, permit the accurate evaluation of job performance on the basis of objective criteria (4302 (b)(1))." Both sample organizations met their statutory requirements. The sample managers received ratings and performance feedback in October 1981 and again in October 1982; the sample employees received their first performance ratings in October 1982. Managers and employees were rated on objectives, and these ratings were summarized as an overall rating through approved formulae. For managers these ratings directly determined merit pay awards, but for employees standard pay increases were given to all those with satisfactory or better performance. The new system required performance-based objectives, a summary rating, and face-to-face feedback with the subordinate signing the appraisal form and required that these ratings be the basis for personnel decisions.

Procedure

The study used a time-series, repeated measures quasi-experimental design (Cook & Campbell, 1979). The present sample is composed of a panel of respondents who completed attitude questionnaires at the following series of time periods across a total of 30 months.

Time 1—Pre-ratings period. Attitude questionnaires were distributed to respondents by members of the research team at three points in time before either managers or employees received formal ratings. If a respondent provided complete attitude data in June 1980, December 1980, and June 1981, or any two of these, the mean of the responses is used as the Time 1 value. Whenever more than one Time 1 response was available the responses were averaged rather than the random selection of one questionnaire administration, so that all available data could be used to provide a more stable measure of pre-ratings attitudes. If the respondent completed the questionnaire only once during this period, the single value is used.

Time 2—First ratings feedback for managers and pre-ratings for employees. The fourth survey was conducted in December 1981, 2 months after the first ratings feedback was received by managers in October 1981, but ten months before the employees received their ratings.

Time 3—Second ratings feedback for managers and first ratings feedback for employees. The final survey was conducted in December 1982, 2 months after the second year-end ratings feedback to managers and the first ratings feedback to employees.

Respondents were fully informed that this was an Office of Personnel Management funded evaluation of the act, that their responses would be completely anonymous, and they were given brief "letter format" feedback after each questionnaire session. Respondents were randomly selected, and the average response rate was 81%.

Measures

The first independent variable of self-reported rating was taken from a question asked only at the final questionnaire administration (i.e., Time 3). Respondents were asked "What was your annual performance appraisal rating last year?" Note that this would be the second rating for managers and the first rating for employees. Those indicating that they knew what their rating was (94% of respondents) had the following choices: "outstanding," "highly successful/highly satisfactory/excellent/exceeds acceptable," "successful/satisfactory/fully satisfactory/acceptable," and "unsatisfactory/unacceptable." Those in the first two categories were grouped together as "high" (64% of sample), the next category is called "satisfactory self-reported rating" (36% of sample), and there were no respondents reporting unsatisfactory performance.

The second independent variable, actual rating, is available only for research center managers and employees. A comparison of actual and self-reported ratings for center employees and managers is reported in the Analysis section.

The dependent variables consist of three attitude scales, representing the respondents' (a) perceptions of appraisal system operation and organizational impact and (b) commitment to the organization.

1. *Perceptions of appraisal system operation and organizational impact.* These are reflected in Operation of System and Effects of Performance Appraisal on Organization scales. The questionnaires contained statements about present appraisal practices to which respondents indicated their agreement/disagreement on 7-point scales. These were factor analyzed separately for managers and employees, but two identical scales could be constructed for both groups by using a principal components factor analysis procedure with a varimax rotation of factors. Operation of the System is composed of six items, including, for example, "my supervisor and I agree on what 'good performance' on my job means," "I am *not* sure what standards have been used to evaluate my performance (reverse score)," and "the standards used to evaluate my performance have been fair and objective." The scale has an $\alpha = .86$ for both managers and

Table 1
Means, Standard Deviations, and Intercorrelations for Time 3

Variable	Managers			Employees						
	<i>M</i>	<i>SD</i>	1	2	3	<i>M</i>	<i>SD</i>	1	2	3
1. Operation of system	5.14	1.07	—			5.14	1.14	—		
2. Effects of performance appraisal on organization	3.71	1.34	.43***	—		3.92	1.26	.48***	—	
3. Organizational commitment	5.00	0.87	.35***	.25**	—	4.78	1.21	.47***	.40***	—

Note. Manager $n = 101$; Employee $n = 348$.

* $p < .05$. ** $p < .01$. *** $p < .001$.

employees. "Effects of Performance Appraisal on Organization" consists of four items, for example, "the present performance appraisal system contributes to overall agency effectiveness," and "all in all, I feel the current performance appraisal process is effective;" it has an $\alpha = .85$ for managers and $\alpha = .82$ for employees. The means, standard deviations, and intercorrelations among the dependent variables for managers and employees appear in Table 1.

2. *Commitment to the organization.* This scale constitutes the short form of Mowday, Steers, and Porter's (1979) Organization Commitment Scale. It consists of nine statements to which respondents are asked to indicate the strength of their agreement or disagreement on a 7-point scale. Sample statements include "I really care about the fate of this organization" and "I am willing to put in a great deal of effort beyond that normally expected in order to help this organization be successful." Mowday et al. (1979) report reliability coefficients ranging from .82 to .93 in eight separate samples of such diverse employees as psychiatric technicians, auto company managers, and bank employees. Organization Commitment has an alpha equal to .89 for managers and .92 for employees in the present sample.

Analysis

These data were analyzed as two-factor experiments with repeated measures on one factor (Winer, 1971, pp. 514-524) using the BMDP statistical package 15.3P4V (Dixon et al., 1981). The factors are Rating at two levels: High and Satisfactory, and Time at three levels (with the same manager and employee panels of respondents queried at all three time periods). As Winer (1971) notes, differences between those receiving High and Satisfactory ratings may include differences between these groups not directly associated with the factor itself; for example, some third variable such as "internal locus of control" may influence both Rating and attitudes. Yet the main effect of Time, as well as the interaction effect, will be free of such confounding. That is, a simple finding that those receiving relatively high ratings consistently have more positive attitudes cannot be interpreted as Rating causing the effect. However, if there is a significant drop in attitudes after the introduction of performance appraisal only for those receiving satisfactory ratings, this finding can be interpreted as resulting from receiving different ratings.

Before reporting the results of the tests of the hypotheses, self-reported and actual ratings for center managers and employees can be compared. There was a high correspondence between the two ratings $\chi^2 (4, N = 124) = 171.52, p < \text{than } 0.001$, for employees; and $\chi^2 (4, N = 33) = 48.58, p < 0.001$, for managers. However, an examination of the direction of the misreports indicates a moderate upward bias that would not be expected if errors were random. Of the 19% misreporting employees, 15% reported too high while only 4% reported too low; for the managers, 3% reported too high, with none reporting too low. We might speculate that the managers' error rate was so small because they rate their own subordinates and have a clearer understanding of the meaning of the rating categories. The upward bias could be the result of ego-defensive

bias as well as honest misinterpretation of a supervisor's "softening" of the rating feedback (e.g., "I had to give everyone a three, but I think your performance is excellent . . ."). Whether or not an average misreport of slightly more than one out of eight respondents should be regarded as a nontrivial proportion of misreporting, it seems clear that if someone leaves a feedback session perceiving themselves (for whatever reason) to be a high performer, the actual satisfactory ratings in their personnel file will not create cognitive dissonance. However, both actual and self-report ratings data are presented in the tables, and points of divergence are noted in the text.

The self-reported data are combined to make hypothesis testing results clearer to the reader. Separate analyses for each organization did not result in a substantial modification of the combined findings.

Results

Hypothesis 1

It was expected that the receipt of a satisfactory appraisal rating would lead recipients to disparage the operation of the appraisal system and its effectiveness for the organization. As indicated in Tables 2 and 3, there is mixed support for Hypothesis 1.

Regarding operation of the system, receipt of a satisfactory rating between Time 1 and Time 2 resulted in a statistically significant drop in attitudes toward the system, with no change in attitudes for highly rated managers. Yet, both groups of managers viewed the system more positively after the second year of operation (attitudes increased between Time 2 and Time 3). These combined results are clarified by an examination of the results for the separate organizations. At the research center relatively highly rated managers were significantly more positive about the appraisal system over all three time periods, whereas the satisfactory managers' attitudes dropped between Time 1 and Time 2 (consistent with Hypothesis 1); however, both groups reported sharply more positive attitudes after their second feedback session in December 1982 (5.05 vs. 5.61, $p < .01$ for the Highs category; 4.57 vs. 4.97, $p < .05$ for the Satisfactory category). Alternatively, the results for the engineering station managers were consistent with Hypothesis 1: There was a significant drop in attitudes toward the appraisal system after receipt of a satisfactory rating (between Times 1 and 2), and attitudes remained significantly lower a year later (5.16 vs. 4.01 vs. 4.41; difference between Time 1 and Time 2, $p < .01$; difference between Time 1 and Time 3, $p = .01$). In contrast, the more highly rated engineering station managers reported more positive attitudes toward the operation of the system after receipt of their first appraisal. It appears that Hypothesis 1 found strong support in the reports of engineering

station managers, but the drop in attitudes for satisfactory research center managers was reversed after the second year of system operation. (The sharp improvement in the attitudes of all research center managers and employees toward the system after the second round is addressed in the Discussion section.)

For employees, we found an overall improvement in attitudes toward the system across time. Yet the change appears to be confined solely to the reports of employees at the research center (4.96 vs. 5.35, $p = .001$), paralleling the attitude increase of their managers. There were no changes over time or differences between the High and Satisfactory respondents for engineering station employees.

Finally, with respect to reports of the effects of performance appraisal on the organization, the pattern of results is similar to the other tests of Hypothesis 1. For the managers there is an overall nonsignificant trend consistent with Hypothesis 1 and a significant effect for Time. That is, both highly and satisfactorily

rated managers reported significantly more negative views of the effectiveness of appraisal after initial appraisal feedback, but a significant improvement after the second year of operation. Again, this sharp improvement between Time 2 and 3 appeared only at the research center. The results for the engineering station managers are exactly as predicted in Hypothesis 1: There was a sharp drop in reports of appraisal effectiveness after receipt of a satisfactory rating, and they remained significantly lower a year later (4.22 vs. 3.01 vs. 2.90; difference between Time 1 and 2, $p < .05$ difference between Time 1 and 3, $p < .05$), with no change for the highly rated. There were no significant differences between highly or satisfactorily rated employees, after the introduction of formal appraisals in either the combined or individual samples.

To summarize for Hypothesis 1, the results for the engineering station managers are consistent with the hypothesis on both measures, but for the research center managers the initial drop

Table 2
Analysis of Variance

Variable	Source	Managers			Employees		
		MS	df	F	MS	df	F
Operation of system	Rating (A)	6.72 (7.88)	1 (1)	2.88* (3.56*)	27.51 (5.67)	1 (1)	10.21** (2.21)
	Error	2.33 (2.21)	82 (30)		2.70 (2.56)	398 (116)	
	Time (B)	1.77 (3.47)	2 (2)	2.46* (6.16*)	5.03 (6.85)	2 (2)	8.28** (10.60**)
	A × B	5.55 (1.37)	2 (2)	7.72** (2.43*)	1.00 (1.21)	2 (2)	1.64 (1.88)
	Error	0.72 (0.56)	164 (60)		0.61 (0.65)	796 (232)	
Effects of performance appraisal on organization	Rating (A)	4.95 (16.16)	1 (1)	1.76 (6.39**)	0.10 (3.32)	1 (1)	0.03 (1.06)
	Error	2.83 (2.53)	82 (30)		3.09 (3.12)	400 (118)	
	Time (B)	10.23 (3.93)	2 (2)	11.66** (5.00**)	0.98 (2.84)	2 (2)	1.25 (3.20**)
	A × B	2.30 (1.36)	2 (2)	2.62* (1.73)	0.19 (1.65)	2 (2)	0.24 (1.86)
	Error	0.88 (0.79)	164 (60)		0.78 (0.89)	800 (236)	
Organizational commitment	Rating (A)	0.61 (4.31)	1 (1)	0.34 (2.75*)	21.12 (10.76)	1 (1)	6.56** (4.49**)
	Error	1.83 (1.57)	82 (30)		3.22 (2.40)	401 (120)	
	Time (B)	0.33 (0.14)	2 (20)	1.43 (0.56)	0.76 (0.41)	2 (2)	2.45* (1.41)
	A × B	1.06 (0.76)	2 (2)	4.60** (3.07**)	1.34 (0.39)	2 (2)	4.33** (1.36)
	Error	0.23 (0.25)	164 (60)		0.31 (0.29)	802 (240)	

Note. Numbers in parentheses are statistics for Actual Research Center Ratings.
* $p \leq .10$. ** $p < .05$.

Table 3
Means of the Dependent Variables

Variable	Rating	Managers			Employees		
		Time 1 ^a	Time 2	Time 3	Time 1 ^a	Time 2 ^a	Time 3
Operation of system	High	4.71 (4.89)	5.08 (4.89)	5.27 (5.63)	4.99 (4.86)	5.12 (4.90)	5.31 (5.47)
	Satisfactory	4.98 (4.79)	4.33 (4.15)	4.74 (4.74)	4.73 (4.69)	4.87 (4.80)	4.87 (4.98)
Effects of performance appraisal on organization	High	3.94 (4.16)	3.54 (3.90)	3.78 (4.21)	3.83 (3.74)	3.87 (3.54)	3.94 (3.89)
	Satisfactory	4.03 (3.75)	2.97 (2.66)	3.38 (3.38)	3.80 (3.68)	3.92 (3.96)	3.88 (4.10)
Organization commitment	High	4.96 (5.09)	5.05 (5.25)	5.06 (5.27)	4.95 (5.27)	4.90 (5.14)	4.95 (5.26)
	Satisfactory	5.12 (5.02)	4.77 (4.60)	4.87 (4.71)	4.71 (4.97)	4.72 (4.89)	4.54 (4.79)

Note. Numbers in parentheses are statistics for Actual Research Center Ratings.

^a Pre-ratings period.

in the attitudes of satisfactory managers is reversed after the second feedback session. Among employees in both organizations there was, by and large, no support for Hypothesis 1.

Hypothesis 2

It was hypothesized that those receiving satisfactory ratings would have more negative attitudes about the organization, whereas the feelings of those receiving relatively high ratings would remain unchanged. The analysis of variance tests of this hypothesis and the mean scores for the variables appear in Tables 2 and 3, respectively.

As can be seen in Table 2, the second hypothesis was supported. The interaction terms are significant for both managers and employees (except for the actual ratings of research center employees, which do not reach statistical significance). In Table 3 it can be seen that the organizational commitment of managers rated satisfactory dropped significantly after their first appraisal feedback (between Time 1 and Time 2) and remained significantly lower a year later, whereas the organizational commitment of highly rated managers neither increased nor decreased during this period of time. Furthermore, the organizational commitment of satisfactory employees dropped significantly only after their appraisal feedback—between Times 2 and 3—with the commitment of highly rated employees remaining stable between June 1980 and December 1982. The separate results for each of the two organizations show the same patterns, except for the actual ratings of research center employees noted above.

These results provide strong support for the argument that feedback that one is a satisfactory but below average performer results in a significant drop in these employees' organizational commitment. Only those receiving satisfactory ratings reported a change in their commitment to the organization; furthermore, this change occurred within 2 months of receiving ratings feedback—in December 1981 for managers and December 1982 for employees—and remained at this reduced level a year later (for

managers, the only group for which two postratings measures are available).

Discussion

The results only partially support Hypothesis 1 but are consistent with the predictions of Hypothesis 2. Hypothesis 1 proposed that the receipt of a below-average satisfactory rating would lead to increased negative views of the operation of the appraisal system and its effects on the organization. This prediction was not supported for the nonmanagement employees. In fact, all research center employees showed significant improvements in attitudes toward how well the appraisal system operates between Times 2 and 3. One reason for the failure of Hypothesis 1 to be confirmed for this group of nonmanagement employees could be simply the fact that perceptions of how well the system operates and its effects on the organization are relatively "distant" in a psychological sense and not very important one way or the other to those who are neither supervisors nor personnel specialists. A related reason may be the fact that the employees' ratings were not to be used for pay increases, but a portion of the managers' pay increases were based on their performance ratings. Under these circumstances, it would probably be easier for employees to reduce experienced dissonance by minimizing the importance of their appraisal ratings. However, because Hypothesis 1 was generally confirmed for the managers, the salience of attributes of the performance appraisal system may be greater for those with supervisory responsibilities and under merit pay systems, and hence their attitudes toward the appraisal system would be more likely to be affected by the receipt of "unexpected" satisfactory ratings.

The other trend in the findings that did not consistently support the predictions of Hypothesis 1 concerned the research center subsample. For those who worked in this agency, attitudes toward the appraisal system and towards its perceived impact on the organization increased between Time 2 and Time 3 regardless

of whether the individuals received relatively high or satisfactory ratings. Although the reasons for this are not entirely clear, it should be noted that the management of the research center devoted intensive personal effort between Time 2 and Time 3 to improving the operation of the appraisal system. This appears to have had the effect of making most individuals, managers and nonmanagers, more positive about the appraisal system regardless of the rating they received (see Landy et al., 1980). It may be that, in effect, the organization's actions to improve the appraisal system inoculated individuals from disparaging the system, even if they received a rating lower than they expected or felt they deserved. It should be noted, however, as will be discussed below, that efforts to improve and perfect the appraisal system did not prevent a subsequent drop in organizational commitment on the part of those who received satisfactory ratings. It may be that the extensive organizational attention to improving the research center's appraisal system simply removed it as a likely target for blame by those who received feedback that they were satisfactory but did not have an impact on experienced dissonance or on resultant blame of another target, namely, the organization.

The major finding of the present study, consistent with the predictions of Hypothesis 2, is that receipt of relatively low ratings caused a distinct and significant drop in attitudes toward the organization within 2 months of feedback of the appraisal results. This occurred in both sites and for both management and nonmanagement employees and persisted a year later. Whether this type of impact should be of concern to organizations, assuming it would be replicated in other types of samples, would depend on how particular organizations view those who are performing adequately but are not superior or clearly above average. In certain organizations (e.g., major public accounting firms), where only the most outstanding performers are retained after a few trial years, the drop in positive attitudes toward the organization probably would not be a cause for concern. For other organizations, however, where turnover of adequate performers is costly or where (as in a research and development type of organization) all such members are considered important contributors, such reactions from those who receive relatively—but not necessarily absolute—low ratings may be a greater problem.

Before concluding, several limitations of the study should be noted. First, the data on individual performance at the engineering station (which provided results most consistent with the hypotheses) consisted of self-reports of ratings received from supervisors, rather than the actual records of those ratings. The comparative data from the research center indicates that, although respondents are substantially correct in their reports, there is a small upward bias in self-reports. However, as noted above, it is the respondent's perceived feedback that should best predict their reactions.

A second limitation of the design was the fact that no items were included in the questionnaire regarding a major factor in the performance ratings environment, namely, the supervisor. The reason for this is simply that the chief aim of the larger project, of which the present study is only a small segment, was to assess the organizational impacts of the Act. Hence, at the time the questionnaires were designed prior the start of the overall investigation, items were not included that focused on the individual's reaction to the supervisor. Obviously, such data would have been useful and pertinent to the question of the impact of

received ratings had items of this type been included as part of an already-lengthy questionnaire.

The third limitation related to the nature of the particular sample of respondents. In this sample, a rating of "satisfactory" may be more psychologically negative than would perhaps be the case with some other types of samples. The dominant employee groups in both samples were highly educated scientists and engineers. Many of these individuals reported working for the federal government, rather than in private organizations, because they were able to do "state of the art" work in these agencies. Clearly, these employees, as would be the case for other high status professionals, found the feedback that they were merely satisfactory to be negative feedback, indeed. Yet, this reaction may not be as severe among members of other kinds of occupations.

Fourth, the study was restricted to attitudinal variables, and there were no tests of their direct effects on behaviors. We know that lowered organizational commitment is associated with reduced attendance and increased turnover in other samples (Mowday, Porter, and Steers, 1982) but have no evidence of its impact on these research center and engineering station employees. Thompson and Dalton (1970) argued that the "demoralization" resulting from such feedback would reduce individual's performance because they would see little likelihood that increased effort would result in a major movement in their relative ranking. However, Mowday et al. (1982) found no significant association between commitment and individual performance in their extensive review.

Finally, the present study is limited in its focus only on the first few months (for nonmanagement employees) and first 14 months (for managers) after the implementation of a new formal appraisal system. The feedback that one is "merely average" no doubt loses its shock value upon repetition year after year. The present study did provide evidence that the decrement in attitudes toward the appraisal system can be temporary (for research center employees) and it may be that other nonperformance feedback issues become a more salient factor in one's organizational commitment over time.

Nevertheless, it appears that Thompson and Dalton's (1970) and Meyer's (1975) concerns about the possible unintended negative consequences of "overly precise," implicitly comparative, appraisal systems may be well founded, given the results found in this study regarding the sharp and immediate drop in organizational commitment on the part of those receiving satisfactory ratings. It seems clear that for many people self-perceptions of their organizational or work performance are closely aligned with their feelings of self-esteem, and thus they want to believe that they are making important contributions. Any appraisal system, then, that provides data, whether explicit or implicit, on how one ranks compared to one's peers is likely to generate some loss in positive feelings on the part of those who are not (as is the case in any ranking system) in the upper part of the distribution. Furthermore, this study demonstrates that such reactions are not confined only to those with clearly unsatisfactory performance but may extend to most of those below the upper rankings. This would suggest that organizations need to consider carefully how their appraisal systems affect not only the attitudes and performance of those ranked at the top, but also those "solid citizens" who are performing at acceptable, but not outstanding levels.

This is not intended to imply that formal performance appraisals should be abandoned. This problem is primarily one of perception and interpretation, and there are several practical steps that personnel specialists can take to minimize these negative attitudinal effects. For example, performance appraisal systems can be examined for any unintended or unnecessary (for the uses of the system) ranking of employees. If the ratings will not be used for any comparative purposes (e.g., pay allocation from a fixed pool, layoff decisions), the potential attitudinal costs of implicit or explicit employee performance rankings would suggest that the system be modified. Furthermore, personnel specialists can do much to anticipate the potential negative attitudinal consequences through training supervisors. Supervisors can be helped to anticipate and more effectively manage the possible negative reactions of employees to such feedback.

In conclusion, the results of the present study suggest that perhaps less attention should be paid to refinements of the psychometric details of rating instruments and relatively more attention paid to the organizational behavior impacts of performance ratings feedback. Performance appraisal takes place in a complex social system, and feedback concerning relative performance is an important signal to employees about how their organizations value them. Thus, any system that drifts into a pattern (and an associated way of thinking) of separating and identifying the "stars" from the "also rans" is likely to have effects that may not be intended or organizationally desired, regardless of the technical quality of the measurement system.

References

- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Boston: Houghton Mifflin.
- Carroll, S. J., & Schneider, C. E. (1982). *Performance appraisal and review systems*. Glenview, IL: Scott, Foresman.
- Dixon, W. J., Brown, M. B., Engleman, L., Frane, J. W., Hill, M. A., Jennrich, R. I., & Toporek, J. D. (1981). *BMDP Statistical Software*. Berkeley: University of California Press.
- Feldman, J. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology*, *66*, 127-148.
- Festinger, L. (1961). The psychological effects of insufficient rewards. *American Psychologist*, *16*, 1-11.
- Ilgen, D. R., Fisher, C. D., & Taylor, M. S. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, *64*, 349-371.
- Ilgen, D. R., Peterson, R. B., Martin, B. A., & Boeschen, D. A. (1981). Supervisor and subordinate reactions to performance appraisal sessions. *Organization Behavior and Human Performance*, *28*, 311-330.
- Landy, F. J., Barnes-Farrell, J., & Cleveland, J. N. (1980). Perceived fairness and accuracy of performance evaluation: A follow-up. *Journal of Applied Psychology*, *65*, 355-356.
- Landy, F. J., & Farr, J. L. (1983). *The measurement of work performance*. New York: Academic Press.
- Meyer, H. H. (1975). The pay-for-performance dilemma. *Organizational Dynamics*, *3*, 39-50.
- Meyer, H. H., Kay, E., & French, J. R. P. (1965). Split roles in performance appraisal. *Harvard Business Review*, *43* (1), 123-129.
- Mowday, R. T. (1983). Beliefs about the causes of behavior: The motivational implications of attribution processes. In R. M. Steers & L. W. Porter (Eds.), *Motivation and work behavior* (3rd ed.). New York: McGraw-Hill.
- Mowday, R. T., Porter, L. W., & Steers, R. M. (1982). *Employee-organization linkages*. New York: Academic Press.
- Mowday, R. T., Steers, R. M., & Porter, L. W. (1979). The measurement of organizational commitment. *Journal of Vocational Behavior*, *14*, 224-247.
- De Nisi, A. S., Randolph, W. A., & Blencoe, A. (1980, November). Peer evaluations: Causes and consequences. *Proceedings of the 12th Annual Meeting of the American Institute of Decision Sciences* (pp. 420-431). Las Vegas, NV.
- Parker, J. W., Taylor, E. K., Barrett, R. S., & Martens, L. (1959). Rating scale content III: Relationship between supervisory- and self-ratings. *Personnel Psychology*, *12*, 49-63.
- Perry, J. L., & Porter, L. W. (1981). *Organizational assessments of the Civil Service Reform Act of 1978*. Washington, D.C.: U.S. Office of Personnel Management.
- Smircich, L., & Chesser, R. J. (1981). Superiors' and subordinates' perceptions of performance: Beyond disagreement. *Academy of Management Journal*, *24*, 198-205.
- Taylor, M. S., Fisher, C. D., & Ilgen, D. R. (1984). Individuals' reactions to performance feedback in organizations: A control theory perspective. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resources management* (pp. 81-124). Greenwich, CN: JAI Press.
- Thompson, P. H., & Dalton, G. W. (1970). Performance appraisal: Managers beware. *Harvard Business Review*, *48*, 149-157.
- Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.

Received July 1, 1985

Revision received September 11, 1985 ■