Determinants of Perceived Fairness of Performance Evaluations

Jerald Greenberg
Faculty of Management and Human Resources, Ohio State University

Middle managers from three organizational samples responded to an open-ended questionnaire in which they described the determinants of particularly fair or unfair performance appraisals. By Q-sort procedure, the responses were categorized and combined to yield seven distinct determinants of fairness in performance evaluations. Ratings of the perceived importance of these determinants were factor analyzed, revealing two distinct factors—procedural determinants and distributive determinants. The implications of the reported determinants are discussed with respect to existing research and theory on justice in organizations.

What makes a performance evaluation perceived as being fair? In view of the potential importance of fair evaluations in determining workers' acceptance of appraisal systems (Dipboye & de Pontbraind, 1981; Lawler, 1967), this is an important question. Two approaches to the answer may be identified.

Traditional views of justice in organizational settings, such as equity theory (Adams, 1965), have focused on the relative ratio of a worker's outcomes of his or her inputs to some standard of comparison as the basis for assessing the fairness of a relationship (for a review, see Greenberg, 1982). This distributive justice perspective focuses on the fairness of the evaluations received relative to the work performed. Workers may, in fact, assess the fairness of their appraisals by comparing the relative ratings they received in return for the work they contributed (Greenberg, in press-a).

Additional evidence suggests, however, that beliefs about fair performance evaluations may also be based on the procedures by which the evaluations are determined apart from the ratings received (see Greenberg, in press-b). This procedural justice perspective focuses on the fairness of the evaluation procedures used to determine the ratings. For example, in organizational questionnaire studies, Landy and his associates (Landy, Barnes, & Murphy, 1978; Landy, Barnes-Farrell, & Cleveland, 1980) found that the fairness of performance evaluations was related to several process variables (e.g., the opportunity to express feelings when evaluated) regardless of the rating outcomes themselves. Several recent laboratory studies have also shown the importance of procedural variables on the perceived fairness of performance appraisal systems (Greenberg, in press-c; Kanfer, Sawyer, Earley, & Lind, in press).

Although several correlates of perceived fairness in performance appraisals have been revealed by Landy et al. (1978), their study was designed to examine the influence of process variables and, as such, paid little attention to the potential contribution of distributive factors as determinants of appraisal fairness. Moreover, by using a closed-format questionnaire in which the potential correlates of fairness were identified a priori (although on no apparent theoretical basis), it was not possible for subjects to identify any other aspects of the appraisal process that may have contributed to their perception of its fairness. To avoid these limitations, the present study used an open-ended questionnaire method as the basis for identifying the determinants of perceived fair performance appraisals. This practice follows in the tradition of several recent attempts to identify perceived fair managerial practices (e.g., Bies & Moag, in press; Sheppard & Lewicki, in press).

Method

Sample

The research sample consisted of a total of 217 middle managers employed in three different industrial groups: cable-TV companies located throughout the United States (n = 95), wholesale pharmaceutical distributing companies located primarily in the midwest (n = 80), and credit unions in the state of Ohio (n = 42). Demographically, the overall sample was 78% male, 94% white, and 88% college educated, with an average tenure with their present organizations of 4.5 years. Virtually all of the sample claimed to have had at least 30 experiences in receiving and 30 experiences in giving formal performance evaluations.

Procedure and Analytic Technique

The subjects in all three subsamples took part in the study in conjunction with their participation in management training seminars. Separate subsets of the overall sample were used in each of the four phases of the study. Each subset contained a roughly proportional number of subjects from each of the three subgroups. The division into subsets was made on the basis of the availability of groups at various times in their training schedules, although the scheduling of participants to these events was decided at random.

In the response generation phase of the study, 56 participants were asked to think of an incident in which they received either a particularly fair or unfair performance evaluation on their job. They were then asked to write down the one most important factor that made it so fair or unfair. These responses were paraphrased and abstracted into simple statements and reworded in the fair direction by the experimenter. The phrases were then typed on index cards.

1 Because all three samples were very similar demographically and identical results were obtained for each sample separately, no distinction is made between the three subsamples used in this article.
The categorization phase consisted of giving the cards to a group of 40 participants from the same sample who were asked to sort them into similar groupings following the unstructured Q-sort technique (Stephenson, 1953). That is, no predetermined number of categories was specified. To generate a homogeneous composite of these sortings that reduced the number of idiosyncratic statements, the criterion of retaining statement clusters that showed at least 75% common overlap was used. That is, for a response category to be defined, a set of two or more statements had to be grouped together by at least 75% of the subjects. Given that no a priori number of categories was imposed (the obtained range was 3 to 13), this level of agreement is considered to be a very high one in Q-sort research (Stephenson, 1953). This procedure reduced the 56 original statements (1 from each of 56 respondents) to 18 statements falling into seven different categories (containing a range of from 2 to 4 statements each). The within-category rate of interrater agreement ranged from 75% (30 of 40 subjects put the same group of statements in one category) to 100%, with an average within-category agreement rate of 88%.

In the cross-validation phase of the study, 46 additional subjects categorized the 18 statements back into the seven categories. Because of the earlier elimination of idiosyncratic statements, subjects were highly successful in performing these ratings, with a hit rate of 98.6% (of 828, only 12 were misclassified).

In the importance-rating phase of the study, another group of 75 subjects from the same sample were asked to rate the importance of each of the seven categories as determinants of fair evaluations. A 9-point bipolar scale with endpoints labeled not at all important (1) and extremely important (9) was used.

Results

Responses to the seven questionnaire items were factor analyzed using the principal-factors technique and a varimax rotation. The first two factors accounted for 94.7% of the variance, and a third added less than 1%. The factor loadings are displayed in Table 1.

The five items that loaded highly on Factor 1, but not on Factor 2, clearly appear to reflect procedural dimensions of fairness and are so labeled in Table 1. In addition, two items that loaded highly on Factor 2, but not on Factor 1, appeared to reflect distributive dimensions of fairness. It should be noted that although the observed pattern of factor loadings is clear, the relatively small sample size on which the ratings were based (N = 75) makes these loadings more unstable than would be desirable.

The mean importance ratings of each of the statements also appear in Table 1. A one-way repeated-measures analysis of variance comparing the importance of the distributive factors and the procedural factors revealed no significant difference, F(1, 449) = 1.24. Within-subjects comparisons between the five procedural factors revealed no overall difference in their perceived importance, F(4, 296) < 1. Similarly, the ratings of the two distributive factors were not rated as being significantly different from each other, F(1, 74) < 1.

Discussion

The present results provide empirical support for several theoretical conceptualizations of procedural justice that have postulated the importance of various determinants of fair procedures. For example, two of the factors identified in the present study, ability to challenge/rebut evaluations and consistent application of standards, correspond closely with Leventhal, Karuza, and Fry's (1980) identification of appeals procedures and consistency of allocation practices as determinants of fair procedures for the distribution of resources. Similarly, the soliciting input factor and the two-way communication factor identified here reflect Thibaut and Walker's (1975) concern for process control (the opportunity to influence the information that will be used to make decisions)—a concept central to their theory of procedural justice. Although support already exists showing the importance of the Leventhal et al. (1980) and the Thibaut and Walker (1975) determinants of procedural justice (for a review, see Folger & Greenberg, 1985), the present findings highlight their potential applicability to performance appraisal contexts.

The present findings also provide strong support to conceptual attempts to expand procedural justice conceptualizations by applying them to organizational settings in general (Greenberg & Folger, 1983) and performance appraisal situations in particular (Greenberg, in press-b). Greenberg's (in press-c) finding that the use of diaries enhances the perceived fairness of performance appraisals corroborates identification of the rater familiarity with ratee's work factor identified in this study inasmuch as the supervisor's keeping of a performance diary may be seen as increasing his or her familiarity with the ratee's performance. The importance of rater familiarity as a determinant of fair appraisals revealed in this study corroborates Landy et al.'s (1978) significant

Table 1

<table>
<thead>
<tr>
<th>Determinant</th>
<th>Factor loading</th>
<th>Importance rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soliciting input prior to evaluation and using it</td>
<td>.46</td>
<td>6.1</td>
</tr>
<tr>
<td>Two-way communication during interview</td>
<td>.67</td>
<td>5.8</td>
</tr>
<tr>
<td>Ability to challenge/rebut evaluation</td>
<td>.66</td>
<td>4.9</td>
</tr>
<tr>
<td>Rater familiarity with ratee's work</td>
<td>.59</td>
<td>5.1</td>
</tr>
<tr>
<td>Consistent application of standards</td>
<td>.55</td>
<td>6.9</td>
</tr>
<tr>
<td>Distributive factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recommendation for salary/promotion based on rating</td>
<td>.25</td>
<td>7.1</td>
</tr>
</tbody>
</table>

Note. N = 75.

2 Support for the identification of these factors was derived using a cross-validation procedure similar to that used for the generation of the items. A group of 56 undergraduate students at a midwestern university reverse-sorted the seven statements into two categories—procedural factors (concerning the fairness of the evaluation procedure) and distributive factors (concerning the fairness of the resulting evaluation). Of the 392 total categorizations, 385 were accurate, resulting in a hit rate of 98.2%, thus supporting the identification of the factors.
correlation between fairness ratings and a questionnaire item tapping supervisor’s knowledge of performance. It is also consistent with the expressed importance of rater familiarity as a determinant of valid performance appraisals (e.g., Ilgen & Barnes-Farrell, 1984).

There are several additional ways in which the present findings corroborate and extend other studies of procedural justice in organizational settings. Most generally, the present findings support the existence of both procedural and distributive determinants of fairness in organizations found in Alexander and Ruderman’s (in press) large-scale survey study. More specifically, several studies inspired by Thibaut and Walker’s (1975) theory have found that procedures giving employees input into the performance appraisal system are seen as being fairer than those that do not (Kanfer et al., in press; Lissak, 1983). The present study complements these findings by showing that workers are aware of the importance of soliciting and using worker input as a precondition of fair appraisals.

The two distributive factors identified in this study very closely correspond to the author’s distributive justice-based analysis of performance appraisals (Greenberg, in press-a). In this connection, support was found for two distinct channels of influence as determinants of distributive justice—the relationship between performance and rating and that between rating and subsequent administrative action. The significant correlation between perceived fairness and the Landy et al. (1978) item “Is action plan related to performance weak?” reflects a similar sensitivity to distributive factors. The finding that distributive factors were rated as being as important as procedural factors as determinants of fairness suggests that researchers and theorists should not allow distributive factors to get lost in the shadow of the recent attention paid to procedural determinants of fairness. Clearly, it would appear that both procedural and distributive factors need to be taken into account in any thorough conceptualization of justice in organizational settings.

In closing, a note of caution is in order with respect to the generalizability of the results. Although three different managerial groups were used in this study, the great similarity between them does not allow us to conclude that the same factors would emerge in different populations, or that they would have the same relative importance. Indeed, the present results, despite their strong fit with existing research and theory, must be considered limited to the population used in the study. Future research is needed to determine the extent to which other factors would emerge as determinants of fairness in different populations.

References


Received June 17, 1985.