The Performance Appraisal Process: A Model and Some Testable Propositions

THOMAS DECOIIS and ANDRE PETIT
Cornell University

A literature-based model of the determinants of the accuracy of performance ratings is presented. The model indicates that the major determinants of accuracy are: (a) rater motivation; (b) rater ability; and (c) availability of appropriate judgmental norms. Several propositions and suggestions for further research are derived from the components of the model.

Performance appraisals are frequently used in organizations as a basis for administrative decisions such as employee promotion, transfer, and allocation of financial rewards; employee development, including identification of training needs and performance feedback; and personnel research (2). A problem that diminishes the utility of appraisals for these purposes is the seemingly impossible task of obtaining accurate appraisals of employee job behavior, a difficulty most often attributed to faults in the rating format used (38), deficiencies in appraisal content (17), rater resistance to judging others (12, 18), and the implications of the specific purpose of appraisal for the rater and the ratee (3).

Commonly recommended means for obtaining accurate appraisals include changes from traditional trait-based rating formats to behaviorally-based formats (9, 11), rater training (4, 30), and work on the criterion itself (17, 27, 37). While each has its value, no single approach has proven to yield satisfactory results. Few efforts to combine solutions or to develop complementary solutions are reported in the performance apprai-

Received 5/19/77, Revised 8/4/77, Accepted 9/26/77.
Received 11/10/77

Our thanks to Professors Lee Dyer and Craig E. Schneider for their helpful comments on an earlier draft of this manuscript.
The literature tends to be disjointed, lacking the cohesion one would expect in such volume.

Systematic efforts are needed to tie the literature together in terms of a comprehensive and systematic model of the performance appraisal process, a model presented here. The dependent variable is a set of ratings on a performance appraisal instrument, the accuracy of which is shown to be affected by several independent variables. The purposes of the model are to: (a) organize relevant literature in terms of existing theory; (b) provide a set of testable propositions bearing on appraisal problems; and (c) increase our understanding and control of the process in organizations.

A Model of the Performance Appraisal Process

Taft's (39) theory of interpersonal judgment is an excellent first approximation of the performance appraisal process. His theory states that the accuracy of one individual's judgment of another individual is determined by: (a) the judge's motivation to evaluate accurately; (b) the availability of the appropriate judgmental norms or standards to the judge; and (c) the evaluative ability of the judge. Similarly, the accuracy of performance appraisals can be viewed as a function of: (a) rater's motivation to appraise accurately; (b) job-relevance of the rating standards used by the rater; and (c) rater's ability to evaluate ratee job behavior.

From this statement of the model, our first and most general proposition follows:

P1: Accurate performance ratings are most likely to occur when the rater is motivated to rate accurately, uses the appropriate rating standards, and has the ability to correctly interpret ratee job behavior in terms of these standards.

This proposition delineates the major components of our performance appraisal model, but it lacks the specificity necessary to explain the process or to provide guidelines for future research and practice. It is necessary to specify the concomitant variables that affect the major components of our model. Figure 1, based on the performance appraisal literature, specifies determinants of appraisal accuracy and possible linkages between them.

Virtually none of the variables included in the model is new; all have appeared many times in the literature. The model's uniqueness stems from their consideration together. Each of the three major components, their respective determinants, and propositional implications are discussed below, always with a focus on the dependent variable, performance appraisal accuracy.

Rater Motivation

The relevance of the concept of motivation to the appraisal process is suggested by the fact that:

- if the (rater) is motivated to make accurate judgments about his (ratees) and if he feels free to be objective, then he has a good chance of achieving his aim, provided he has the requisite abilities and can use the appropriate evaluative standards (39, pp 176-177)

If so, then consideration must be given to what energizes rater behavior, what determines the direction of such behavior, and how this behavior is sustained in the context of a formal performance appraisal program.

Figure 1 indicates six determinants of rater motivation. Without regard to the order of their importance, they are: (a) perceived consequences of accurate appraisal for both rater and ratee; (b) rater perceptions of the adequacy of the performance appraisal instrument used; (c) relevant organizational policies and practices; (d) rating format itself; (e) availability of appropriate standards of performance; and (f) purpose of appraisal. Each of the six determinants their linkages to other determinants, and the attendant propositions are discussed below.
Perceived Consequences of Accurate Appraisal — As Rose notes . . .

[raters] do not always complete appraisal forms when required to do so; they do not always complete every section of the form . . . they do not always acknowledge their authorship, and finally, the content of their entries is often evasive (35, p 2).

Although several explanations could be offered for this reticence, the most plausible one centers on the inextricable linkage between the purpose of appraisal and the consequences of appraisal for the rater and ratee.

For example, appraisals tend to be less accurate when they are obtained for administrative purposes than when they are obtained for purposes of either employee development or personnel research (3, 14, 20). These differences in accuracy have been explained in terms of the relationship between the purposes of appraisal and the likelihood or appropriateness of confidentiality (3), rater role conflict (12), and the possibly negative impact of appraisal results on future ratee outcomes (18).

By way of illustration, performance appraisals obtained for personnel research purposes are typically not available to the ratee and therefore unlikely to affect future rater-ratee interactions. In addition, they are separate from the rater's customary organizational role and have
no impact on the ratee's attainment of valued organizational rewards. In short, ratings obtained under these conditions pose no threat to the status quo of the rater or the ratee. As a result, the rater may be motivated to assign the performance rating(s) he or she feels most accurately reflect true ratee job behavior.

Performance appraisals conducted for purposes of employee development are likely to be more accurate than administrative appraisals but less accurate than appraisals conducted for purposes of personnel research. Employee development may be a dimension of the rater's organizational role and viewed by him or her as in the best long-run interest of the ratee. Consequently, under certain conditions such as performance goal setting (33) and mutual performance problem solving (31), a rater may be motivated to assign accurate performance appraisal scores. To the extent that the rater is required not only to identify ratee development needs but also to be the catalyst for corrective changes in ratee behavior, inaccurate (e.g., lenient) appraisals as a means of avoiding these responsibilities, may result. It is difficult for many raters to constructively criticize ratee job behavior or to provide unambiguous cues for ratee behavioral change (12, 33).

Performance appraisals obtained for administrative decisions are likely to be the least accurate. Analogous to the proverbial forward pass in football, performance appraisals obtained for administrative purposes can lead to three outcomes, two of which are inherently bad; specifically, the ratee is rewarded, not rewarded, or punished. In an organization that attempts to build a perceived linkage between the administration of organizational rewards and job behavior, the results of appraisal will have to be justified to the ratee. Insofar as the rater is required to make these justifications (3), feels uncomfortable doing so (32), or is reluctant to have a negative impact on ratee outcomes (18), or on future rater-ratee interactions (12, 20), he or she will be motivated to assign inaccurate performance ratings.

A reasonable conclusion that can be derived from the above discussion is that most consequences for the rater of assigning accurate performance ratings are negative. Thus, many raters approach the performance appraisal event with reluctance and assign inaccurate ratings simply because there are so few benefits inherent in assigning accurate ratings and so few penalties associated with inaccurate ratings. Hence, our propositions with respect to the motivation to rate accurately are.

**P1A:** Rater motivation to assign accurate performance ratings is higher when the purpose of appraisal is personnel research than when the purpose is either employee development or administration of organizational rewards.

**P1B:** Rater motivation to assign accurate performance ratings is higher when the purpose of appraisal is employee development than when the purpose is administration of organizational rewards.

**P1C:** Regardless of purpose, rater motivation to assign accurate performance ratings is higher when the results of the appraisal are confidential from the ratee.

**P1D:** When the purpose of appraisal is either employee development or administration of organizational rewards and the rater is required to discuss the results of appraisal with the ratee, rater motivation to assign accurate performance ratings is higher when the rater perceives feedback as a legitimate aspect of his or her role.

**P1E:** When feedback of the results of appraisal is required, rater motivation to assign accurate performance ratings is higher if the rater perceives that he or she has the necessary insights into ratee job behavior.

**Perceived Adequacy of the Instrument**

Two characteristics of performance appraisal literature are the large variety of performance ap-
praisal formats it makes available to users and the small number of studies reporting their relevance to the purposes of appraisal. While variety is the spice of life for many phenomena, in the instance of performance appraisal it may be more the result of a serial process of user dissatisfaction than a true need for so many alternatives. Indeed, the literature appears to reflect the following cycle: a particular appraisal technique is first widely praised, then used and rejected, and subsequently replaced by a new more highly praised technique. As a consequence, the literature is a catalogue of performance appraisal formats, each with its assumed advantages, adherents, and critics (10, 36). In the absence of assessed adequacy for the purposes of appraisal, it is difficult rationally to sort out, compare, and choose among available alternatives or to determine if development of a new format is warranted.

Only one study compared completely independent performance appraisal instruments in terms of their adequacy to the purposes of appraisal (15). In this study, raters reported a general preference for a trait-based instrument, while indicating that a numerically anchored behavioral-based instrument was easiest to understand and most relevant for employee development purposes. Hence, our assessment of the impact of the perceived adequacy of an instrument on rater motivation must be left to conjecture.

Logically, a rater’s perception of the adequacy of an instrument should affect the accuracy with which it is used. The most important determinants of this perception are thought to be how easily the rater can understand the instrument (e.g., its content and instructions) and its apparent bearing on a given purpose. If the content of a performance appraisal instrument is based on job analysis, or on some other form of job information, it will have the potential of being both face and content valid and thereby of having a positive impact on rater perceptions of its adequacy for the purpose at hand. Hence, our second set of propositions is:

P2A: The easier a performance appraisal instrument is to understand, the more likely raters are to perceive it as being adequate for the purpose at hand.

P2B: A performance appraisal instrument that is based on systematically gathered job information will be perceived by a rater as more adequate for the purpose of appraisal than an instrument developed on an ad hoc basis.

P2C: A performance appraisal instrument that is based on job information, easy to understand, and perceived as adequate for the purpose of appraisal will lead to increased rater motivation to assign accurate performance ratings.

Rater Ability

As indicated in Figure 1, another major determinant of the accuracy of performance ratings is rater ability, defined as the skill with which a rater interprets ratee job behavior. The determinants of rater ability are: (a) rater’s opportunity to observe ratee job behavior; (b) rater characteristics; (c) rater training; and (d) availability of the appropriate rating standards.

Rater Opportunity to Observe Ratee Job Behavior — Guion points out that “ratings must come from people who are qualified to rate. The primary qualification is first-hand knowledge. A rater can be qualified only on the basis of his own observations of ratee job behavior.” (23, p. 111). Guion’s comment is the basis of performance appraisal — ratings should accurately reflect ratee job behavior. Indeed, the assumption that appraised performance is a sample of observed behavior underlies the performance appraisal literature. For example, developmental procedures for the currently popular behaviorally anchored rating scales format (38) are based on the assumption that performance should be limited to potentially observable job behaviors if accurate performance appraisals are to result (17).
Whitlock (41) describes this relationship between rater observations of behavior and rating accuracy in terms of the psychophysical law, while Gordon (22) suggests that the relationship is moderated by the correctness of the behavior observed. Specifically, Gordon suggests that raters can evaluate correct job behaviors more accurately than incorrect job behaviors. Apparently, raters know when something is done properly but are less sure of their evaluations when behavior is incorrect. Although we have no evidence, it also may be that raters are less willing (motivated) to be candid when observed behavior is incorrect.

Figure 1 also shows that ratee job characteristics affect the opportunity to observe ratee job behavior and, consequently, rater ability. The essence of this linkage is simply that some ratee jobs afford the rater more opportunities to observe ratee behavior than do other jobs. Obviously, a rater has opportunities to observe a ratee's job behavior to the extent that they work in close proximity. Thus, our propositions with respect to observations of job behavior are:

P3A: The more opportunities to observe ratee job behavior the rater takes advantage of, the higher his or her ability to assign accurate performance ratings.

P3B: The more correct the behavior observed by the rater, the higher his or her ability to assign accurate performance ratings.

Rater Characteristics — Our review of the performance appraisal literature suggests several characteristics of the rater that may bear on his or her ability to assign accurate performance appraisal scores. Among the more salient characteristics documented by the literature are the rater's: (a) position in the organizational hierarchy (40); (b) affect toward the ratee (26); (c) perception of how the work should be done (1); and (d) leadership style (28).

Whita and Tirrell (40) and Guion (23) cautioned against the use of raters from distant organizational levels: " . . . an immediate supervi- sor is better qualified to accurately appraise performance than is his superintendent or department manager, they are too far removed from the person being rated" (23, p. 111). The wisdom of this caveat has been empirically verified by several studies (5, 7, 24) which report unsatisfactory interrater agreement on performance ratings assigned by raters from different organizational levels. Raters from different organizational levels may have different perceptions of the implications of the same ratee behavior for job or organizational outcomes or respond to different aspects of the same ratee's behavior (7). Explanations notwithstanding, these results are uniformly indicative of the meaninglessness of appraising performance using raters from different organizational levels.

They do not address the issue of whether there is differential accuracy in the appraisal of performance within organizational levels; given a defined performance domain, we do not know the impact of rater level on performance score accuracy. Nonetheless, it seems that the closer the rater's organizational level is to the ratee's level, the better his or her opportunities to observe ratee behavior. At the same time, the closer the rater's organizational level is to the ratee's level, the more likely it is that his or her interpretation of these observations will be contaminated by factors independent of job performance. One potentially powerful contaminant is the nature of rater-ratee interactions.

Kallejian et al., (26) report that the nature of these interactions influences accuracy of ratings. The direction and nature of this effect is indeterminate; we do not know whether positive interpersonal relations necessarily result in inaccurately high performance ratings. While it is logically appealing to assume that positive interactions lead to lenient ratings and vice versa, it is reasonable to consider factors that might moderate accuracy such as rater perceptions of the correctness of ratee work procedures.

The relationship between a rater's perception of a ratee's work procedures and the magnitude of the ratings assigned to the ratee has
no clear bearing on the accuracy of appraisal results. Conceivably, for example, the ratee's work procedures could exactly correspond to the rater's perceptions of proper procedures and still be ineffective in terms of Campbell's et al. (9) job behavior → job performance → organizational effectiveness paradigm. In this instance, a high performance rating would be inaccurate. Conversely, a ratee could violate rater perceptions of appropriate work procedures and still be effective in terms of organizational outcomes. Thus, rater perceptions of ratee work procedures may have no systematic relationship to the correctness or eventual effectiveness of ratee job behavior. Unless the rater has accurate perceptions of the implications of particular work procedures for task and organizational effectiveness, he or she probably cannot accurately generalize from work process to work outcomes.

Correctness of procedures notwithstanding, an additional contaminant of performance ratings may be the rater's style of management. Koles (28), for example, found that raters who scored high on the Initiating Structure scale of the LOQ (19), were much less likely to assign lenient overall performance ratings than were raters who scored high on the Consideration scale. Whether these results will generalize to other conceptualizations of leadership is not known.

What is known suggests the following propositions with respect to rater characteristics:

P4A: All else being equal, the closer the rater's organizational level is to the ratee's organizational level, the higher the rater's ability to assign accurate performance ratings.

P4B: Raters from different organizational levels sample different aspects of ratee job behavior. Therefore, the higher the correspondence between the content of the rating instrument used and the rater's sample of ratee job behavior the higher the rater's ability to assign accurate performance ratings.

P4C: Rater ability is higher when the rater understands the linkage between ratee job behavior and organizational outcomes than when he or she does not understand the linkages.

P4D: Rater ability is higher when the rater's personal style stresses goal attainment for ratees rather than consideration of ratees as individuals.

Rater Training — One obvious, but not widely used means for improving rater ability is training (8). Latham, Wexley, and Pursell (30) indicate that workshops and group discussions designed to inform raters about the nature and sources of rating errors can be useful for obtaining increased accuracy in performance appraisal. Similarly, Bernardin and Walter (4) report that raters appraised of the meanings of rating errors and familiarized with the scales to be used showed significantly less leniency error and halo effect.

These findings suggest that if rater training is to be used as a means for increasing accuracy of performance ratings, its content probably should stress the importance of performance appraisal to administrative decisions and personal outcomes, illustrate the sources and implications of rater errors, and thoroughly familiarize the raters with the instrument to be used. In addition to stressing relevant content, methods of rater training should approximate the realities of performance appraisal usage, e.g., simulate the difficulties of feeding back negative evaluations to ratees, etc. Perhaps equally as important as content and form of rater training are organizational policies and procedures that communicate to the raters the importance of accurate appraisals to the organization.

The above discussion suggests two propositions with respect to rater training:

P5A: Raters trained in the principles and problems of performance appraisal have higher ability to assign accurate performance ratings than untrained raters.

P5B: Rater training methods that stress active
learning result in higher rater ability to rate performance accurately than training methods that stress passive learning.

**Availability of Appropriate Rating Standards**

The rating standards used by a rater form the basis of his or her conclusions about ratee performance. As shown in Figure 1, these standards have no direct impact on the accuracy of the performance ratings assigned by a rater. Rather, their effect is through their impact on the rater's motivation and ability to rate accurately. Rating standards are viewed as a function of: (a) ratee job characteristics; (b) ratee personal characteristics; (c) appraisal format; and (d) organizational policies and procedures with respect to performance appraisal.

**Ratee Job Characteristics** — The dynamic nature of jobs and job behavior has been well described and documented (21, 25); specifically, jobs change over time, and behavior varies among individuals in the same job. In the context of performance appraisal, the implications of these dynamics seem clear: job behavior that is effective today may be ineffective tomorrow. A manager who focuses her behavior on the acquisition of new facilities during a period of economic growth might be considered effective, while these same behaviors during a period of economic recession might be considered ineffective. In short, the job has changed while the incumbent has not. Further, in the same situation two quite distinct patterns of job behavior may have equal implications in terms of performance effectiveness. For example, two salespersons may be equally effective in terms of long-run sales volume, but obtain this result in entirely different ways. It might be that one salesperson emphasizes customer relations and repeat sales while the other emphasizes high pressure tactics and large one-time sales.

Hence, our propositions with respect to ratee job characteristics and individual dynamics are:

**P6A**: The higher the correspondence between job context and the rating standards used by the rater, the more accurate the performance ratings obtained.

**P6B**: The more that differences in patterns of ratee job behavior are incorporated into the rating standards used by a rater, the more accurate the performance ratings assigned by the rater.

**Ratee Personal Characteristics** — Figure 1 indicates that in addition to ratee job characteristics, personal characteristics of the ratee impact on the rating standards used by the rater. The most often studied ratee personal characteristics are sex and race. Generally, females are rated lower than males (13, 34), even when the rater is a female. Bigoness (6) suggests that this relationship might be moderated by job level, especially when the female ratee is in a traditionally male role. Race has also been shown to affect accuracy of performance appraisals with whites rating whites higher and blacks rating blacks higher (16). Bigoness (6) found that white raters tended to rate black rates higher than white ratees when performance was poor, but similarly when performance was high.

These ambiguous and tentative findings indicate that more research should be conducted before useful conclusions with respect to the impact of ratee sex and race on ratings can be made. In addition, research investigating the impact of other personal characteristics such as ratee age, tenure in position, job level, etc. is needed. What we do know about ratee personal characteristics suggests the following propositions:

**P7A**: Regardless of rater's sex, females tend to be rated lower than males performing similar tasks.

**P7B**: When rater and ratee are of the same race, the ratee will be assigned higher ratings than when rater and ratee are of different races.

While we recognize the several possibly important moderators of the relationship between
sex, race, and performance ratings, we offer no
moderating propositions due to the paucity of
empirical findings and theoretical statements in
the literature.

Appraisal Format — The format of appraisal
affects the rating standards used by the rater and
has implications for rater training and rater per-
ceptions of adequacy of the instrument used (see
Figure 1). Our concept of appraisal format in-
cludes not only instrumentation, but appraisal
content and developmental procedures as well.
The instrument used to record rater perceptions
of ratee job behavior is viewed as the operational
definition of the appraisal format. Hence, our dis-
cussion of appraisal format focuses on the impli-
cations of the content of appraisal and the de-
velopmental procedures used on the accuracy of
performance ratings.

The content issue is the phrase used to de-
scribe a modest controversy over the emphasis
of performance appraisal. On one side are those
who feel that the basis of performance appraisal
should be the personal traits of the ratee while
the other side stresses the potentially observable
job behaviors of the ratee. The proponents of
trait ratings rationalize their stand on the basis
that there is a practical demand for trait content
(27). In contrast, the behavioral view emphasizes
the tentative and ambiguous nature of the link-
age between personal traits and individual be-
behavior (38), suggesting that it is illogical to ap-
praise often esoteric personal traits that are at
least one step removed from actual job behav-
ior (17, Chapter 3). In short, behavioralists view
themselves as unwilling to accept an uncertain
substitute when the real thing is available.

While each position has its own appeal, re-
cent reviews of the empirical literature (c.f., 36)
invariably lead to the conclusion that the con-
tent issue is far from settled. This lack of closure
is apparently due less to the complexities inher-
ent in assessing the implications of differences in
instrument content than to methodological de-
iciencies that limit the relevance of most com-
parative studies to a resolution of this issue. In
most studies the content issue either is not ad-
dressed at all or, if considered, is obfuscated by
comparing interdependent instruments; that is,
instruments where the only difference is in the
nature of the scale line used or the extent to
which the same dimensions are defined.

Figure 1 suggests that comparisons of for-
mats might be more meaningfully made in the
context of the purpose of appraisal. The same
format may result in differentially accurate per-
formance ratings if it is viewed by raters as more
adequate for one purpose than for another. Al-
though tentative, some evidence supporting this
possibility has been reported by DeCotiis, (15)
who compared a trait-based instrument against
two behaviorally-based instruments. He found
that while raters expressed a general preference
for the trait-based instrument without regard to
the purpose of appraisal and no clear preference
among the instruments when the purpose was
performance counseling, they clearly preferred
one of the behavior-based instruments for the
purpose of identifying training needs. Thus, the
content issue in performance appraisal may be
most meaningfully addressed in the context of
the purpose of appraisal.

The literature is even less helpful with re-
spect to the impact of format on rater training.
Specifically, no empirical studies bearing on this
issue were found, but intuitively it seems that
format, especially content, should have an im-
pact on how well and how easily raters learn to
accurately appraise performance. Logically, to
the extent that the content of an appraisal instru-
ment is easily understood, meaningful, and ac-
ceptable to raters, we would expect that their
task of learning how to use it accurately would
be eased.

Format developmental procedures that in-
clude users as sources of content have been
stressed as one of the means for assuring mean-
ingsfulness and rater acceptance and understand-
ing of the instrument used (38). The data suggest
that participative format developmental pro-
dues have yet to bear the fruitful results hoped
for by their advocates (36). Indeed, as DeCotiis
(15) suggests, rater participation in instrument
development may be largely irrelevant. What may be a more important determinant of rater acceptance and understanding of a performance appraisal instrument is the rater's perception of the correspondence between the instrument and the job, i.e., its content validity. If so, then job analysis-based developmental procedures, rather than rater participation per se, would be an obvious means for assuring this perception.

Our proposition with respect to the format of appraisal is as follows:

P8: The higher the correspondence between the content of appraisal and job content, the higher the likelihood that accurate performance ratings will result.

Organizational Policies and Procedures — Performance appraisal policies and procedures have to do with the uses, accountability for, and administration of performance appraisal programs. These policies and procedures are viewed as determinants of the accuracy of performance ratings through their impact on rating standards used by the rater and the rater's perception of the consequences of accurate appraisal (see Figure 1). One desideratum is that these policies and procedures be consistent with the dominant management philosophy of the organization. For example, in a highly structured autocratic organization, one would expect participative appraisal procedures to meet considerable resistance. Essentially, if a rater is to perceive a performance appraisal procedure as providing the appropriate standards, its use must be perceived as consistent with organizational practice.

Organizational policies and procedures also impact on a rater's perceptions of the consequences of rating. It may be possible to tip the balance of consequences in favor of accurate ratings if the rater perceives that he or she will be held accountable for the ratings assigned. Logically "...raters will attach the same degree of importance to their appraisals that they see being attached to them by their superiors" (35, p. 20). In order to obtain accurate performance ratings it may be necessary for performance appraisal programs to be established with visible and continued support from the highest levels of an organization.

Our propositions with respect to performance appraisal policies and procedures are:

P9A: The higher the correspondence between performance appraisal procedures and the dominant organizational philosophy, the more likely performance will be appraised accurately.

P9B: The more likely it is that a rater will be held accountable for the ratings he or she assigns, the more likely the ratings are to be accurate.

Discussion

The model of the performance appraisal process and discussion of its components suggest three conclusions with respect to performance appraisal research: (a) it is voluminous; (b) largely unsystematic; and (c) at least as subject to fads and fashion as any other aspect of personnel research and practice. Perhaps indicative of the viability of these conclusions is continued emphasis in performance appraisal research and publication on the development and advocacy, respectively, of new appraisal formats as though the primary problems in performance appraisal could be solved simply by changes in the format of appraisal. In short, performance appraisal research has focused on instrumentation at the expense of other, perhaps more important, variables. Current research on topics such as rater training, rating simulation, and use-based scale evaluation (15, 30), indicate that researchers are pursuing other promising avenues for future research. What is not evident is that this research is based on a systematic consideration of the performance appraisal process.

Research conducted in the context of a model may more rapidly advance what appears to be a slow-moving stream of research. Such an
approach might provide researchers with insights into the sources of negative findings, suggest potentially more fruitful variables for attention, and provide a means for systematically assessing and re-evaluating what is known about the performance appraisal process. We are not arguing that our model is the sine qua non of such an approach. No doubt, other variables and relationships could be added to the model. What the model does represent is an initial attempt to view performance appraisal in systematic terms. As such, its purpose is to stimulate further research and thought on performance appraisal as a process of interpersonal judgment.

The next logical step in formal development of performance appraisal is the systematic testing of either the propositions presented here or those derived from some other model. The emphasis should be on the empirical development of new propositions and revision of the model based on the evidence. In the final analysis this more coherent approach probably will result in a much larger and more complex model of the performance appraisal process than the one presented here. More importantly, a systematic consideration of the performance appraisal process will provide researchers and practitioners alike with a better understanding and more effective means of controlling the process in organizations.

REFERENCES

11 Cummings, L L, and D P Schwab Performance in Organizations (Glenview, Ill: Scott, Foresman & Company, 1973)
17 Dunnette, M D Personnel Selection and Placement (Belmont, Calif: Brooks/Cole, 1966)
20 Gellerman, S W The Management of Human Resources (Hinsdale, Ill: Dryden Press, 1976)