PERSONNEL PSYCHOLOGY
1988, 41

AN EMPIRICAL INVESTIGATION OF SELF-APPRAISAL-BASED PERFORMANCE EVALUATION

JIING-LIH FARH, JAMES D. WERBEL, ARTHUR G. BEDEIAN
College of Business Administration
Louisiana State University

This study investigated the effectiveness of a self-appraisal-based performance evaluation system (SABPE) that incorporates self-assessment into traditional supervisory evaluation procedures. Its subject sample consisted of 88 faculty members and their chairpersons at a land-grant state university. Results indicated that (1) there was high congruency between self- and chairperson ratings, (2) both ratings had moderate to high levels of criterion-related validity, and (3) both faculty members and chairpersons reported high SABPE acceptance. The implications of these results for future self-appraisal research are discussed.

A fundamental issue in designing any performance appraisal system is determining the source of the necessary information on which to base assessments (Latham & Wexley, 1981). Supervisor, peer, and self are perhaps the most frequently cited information sources in the performance appraisal literature. In general, however, research indicates that there is only a low to moderate correlation among supervisor, peer, and self-appraisals (Landy & Farr, 1980). Appraisals associated with various information sources (i.e., raters) seem to be based on different perspectives on performance (Borman, 1974). Since different information sources have both advantages and disadvantages, no one source can be universally preferred over another.

Arguably, an ideal performance appraisal system would combine information from multiple sources to form an integrated assessment that maximizes the strengths and minimizes the weaknesses of individual information sources (Latham & Wexley, 1981). Although numerous studies have compared the psychometric qualities of appraisal outcomes on the basis of different rater types (i.e., information sources), very few studies have examined how information provided by different raters can be combined to effectively measure ratee performance. The purpose of this study was two-fold: (1) to investigate a self-appraisal-based performance evaluation

An earlier version of this paper was presented at the Academy of Management meeting, New Orleans, 1987.

Correspondence and requests for reprints should be addressed to Jiing-Lih Farh, Department of Management, College of Business Administration, Louisiana State University, Baton Rouge, LA 70803-6312.

(SABPE) system that incorporates self-appraisals into traditional supervisory evaluation procedures and (2) to examine its effectiveness vis-à-vis criterion-related validity and user acceptance.

Self-appraisal

The use of self as a performance information source is an established practice. People are often called upon in their daily lives to provide information about their own behavior. After all, the self is an ever-present observer of one's actions. Skepticism, however, surrounds the use of self-appraisal as a performance assessment method because of the belief that (1) self-appraisals are subject to self-enhancement desires and (2) most people are unable to evaluate themselves objectively or reliably enough to provide accurate information (e.g., Anderson, Warner, & Spencer, 1984; DeNisi & Shaw, 1977; Levine, Flory, & Ash, 1977).

With regard to this first belief, research shows that self-appraisals of performance tend to be more lenient than either supervisory or peer appraisals (Landy & Farr, 1980; Meyer, 1980; Thornton, 1980). However, nonsignificant or even opposite results have been reported (e.g., Heneman, 1974). Moreover, information is available to suggest that the intended purpose of a self-appraisal may influence the degree to which leniency exists. A recent investigation by Farh and Werbel (1986) showed that self-appraisals conducted for the purpose of distributing rewards were more lenient than those conducted for research purposes. Additionally, consistent with the earlier work of Bauman and Dent (1982), Farh and Werbel (1986) found that leniency diminished when individuals were informed that their self-appraisals would be validated against objective records.

In a meta-analysis of studies in which self-appraisals of ability were compared with measures of performance, Mabe and West (1982) found that 15 of 21 studies reported data indicating that people overestimate their abilities. In line with the above cited work, they also found that self-appraisals of ability tended to be more accurate when their accompanying instructions stated either directly or implicitly that they would be compared with independent criteria.

Taken as a whole, available research suggests that when individuals are asked to rate their abilities or performance, they tend to provide inflated evaluations. The degree to which self-appraisals are inflated, however, seemingly depends on prevailing appraisal conditions. Lenient self-appraisals are seemingly most likely to occur when they may lead to some personal gain or when no independent ability or performance measures are available.

With regards to the second belief that people are unable to evaluate themselves accurately, mounting research evidence counters this notion

(e.g., Klimoski & Hayes, 1980). In a recent review of the self-appraisal literature, Shrauger and Osberg (1981) compared the validity of individual self-appraisal with other procedures commonly used in psychological evaluation (e.g., psychological tests, past performance, peer ratings). Their results indicated that self-appraisals are at least as predictive as are other assessment methods with which they have been compared. Furthermore, there is ample evidence indicating that employees are capable of arriving at reasonably realistic self-appraisals when such appraisals are confined to directly observable performance dimensions (e.g., Downs, Farr, & Colbeck, 1978).

Self-appraisal-based Performance Evaluation

Available research thus suggests that self-appraisals are potentially valuable sources of information for performance evaluation purposes. An open research question, however, is how to incorporate self-appraisals into a traditional performance evaluation process. One suggestion is to use self-appraisals as a basis for performance appraisal interviews. For example, Bassett and Meyer (1968) investigated a self-rating appraisal process at General Electric Company (GE) in which only employees completed performance evaluation forms. Subsequent discussion between managers and their subordinates was based solely on the subordinates' self-appraisals. This was compared with GE's traditional supervisor appraisal approach. Results indicated that (1) the self-appraisals were judged more satisfying and constructive by the supervisors than the traditional supervisor-prepared performance interviews, (2) there was less subordinate defensiveness regarding appraisals, (3) discussions based on self-appraisals more often resulted in future superior job performance than did traditional supervisor appraisals, and (4) low-rated subordinates were especially likely to show improvement in performance after a self-review discussion.

In addition, researchers have noted several potential advantages that may result from incorporating self-appraisals into traditional performance appraisal processes (Carroll & Schneier, 1982; Fletcher, 1986; Latham & Wexley, 1981):

1. Self-appraisal-based performance evaluation systems increase communication between raters and ratees regarding job content, performance criteria, and mutual expectations, thus reducing ambiguity in the appraisal process and resolving rater–ratee disagreement.

2. Self-appraisals not only increase ratee participation, they give ratees a greater sense of control over performance evaluations. This increased participation and feeling of control are typically associated with increased satisfaction and acceptance of appraisal results.

3. Self-appraisals often contain less halo error than supervisory ratings and, thus, are more discriminating across performance dimensions. This may occur because ratees are frequently in a privileged position for observing their own job knowledge and performance. Self-appraisals are particularly valuable where ratees work in isolation or possess a rare skill since they likely possess more information about their performance than any other source.

4. Since ratees and raters occupy different roles, a multiple assessment approach generates a larger data base upon which to make performance evaluation decisions. By tapping multiple perspectives, extremely biased ratings can more easily be identified.

Given the benefits that may accrue from incorporating self-appraisals into traditional performance evaluation systems, it is surprising to find that so little research has investigated their use in this manner. One possible explanation for this lack of research is that few employers are willing to allow researchers to step in and "experiment with" their performance evaluation procedures.

The Present Study

In this study. we had a rare chance to observe a self-appraisal-based performance evaluation (SABPE) process following its incorporation into a university-based traditional performance evaluation system. The new process began by requesting that faculty members complete an activity report to document their performance during the preceding 15 months. Faculty members were then asked to evaluate themselves using a rating form based on this report. The activity report, along with the resulting self-rating, was then returned to the appropriate chairperson, who reviewed the self-appraisal using an identical rating form. On a separate sheet, the chairperson provided written evaluative comments. Chairpersons' evaluations were then returned to individual faculty members. If faculty members disagreed with the ratings or comments, they could discuss both the ratings and comments with their chairperson. If disagreements remained, faculty members could prepare a written rebuttal and attach it to their performance evaluation form. These documents were then submitted to the dean of the faculty member's college to be used in making administrative decisions involving such things as raises and promotions.

In achieving its dual purposes, this study addressed three research questions. First, to what extent are self-appraisals congruent with supervisory appraisals in a self-appraisal-based performance evaluation system? This question can be answered by examining (1) correlations between self-ratings and supervisory ratings, (2) mean differences between self-ratings

and supervisory ratings, and (3) relative variability in self-ratings and supervisory ratings. Second, to what extent are final evaluation outcomes (i.e., supervisory evaluations) correlated with objective indicators of performance? Given the focal sample, we examined the extent to which final supervisory evaluations were correlated with objective performance indicators such as number of articles published and number of committee assignments and whether these final evaluations were correlated with student teaching evaluations obtained independently of the appraisal process. Finally, user acceptance was assessed by asking faculty members and their chairpersons to compare the perceived accuracy, fairness, and preference of the SABPE process with traditional supervisory evaluations.

Method

Sample

The subject sample for this study was drawn from across six departments housed in the business college of a large land-grant university. The sample consisted of 88 full-time faculty members, including 24 full professors, 11 associate professors, 36 assistant professors, 15 full-time instructors, and 2 visiting professors.

SABPE

The annual performance appraisal, which formed the basis for this study, was conducted in April and based on all scholarly activities during the previous 15 months. Department chairpersons first distributed performance appraisal forms to each faculty member with directions for their completion. Faculty members were initially requested to complete an activity report on which they were to list and describe their teaching, research, and service activities for the relevant time period. The activities included classes taught, membership on graduate student committees, curricula development, student advisement, refereed and nonrefereed publications, paper presentations, papers under review, funded research, membership on university and college committees, professional organization activities, and honors/awards received. All activities were specific performance outcomes that could be both quantified and cross validated against archival records.[1]

[1] In our interviews with the chairpersons, it was very clear that the items listed in the faculty activity reports were all verifiable. Although the instructions for completing the activity report did not mention independent verification, they did require faculty members to document each activity. For example, for every meeting paper presented, faculty members were asked to list the title of the paper, the name of the meeting in which it was presented, and when and where the meeting was held. This thorough documentation was not limited to research performance.

On the basis of their activity reports, faculty members were asked to evaluate themselves in three areas: instruction, scholarship, and service. Instruction was divided into two subareas: (1) instructional method, including teaching methods/technique, teaching innovation, and curricula development; (2) instructional support, including student advisement, service on graduate committees, and participation with student organizations. Scholarship activity was likewise divided into two subareas: (1) journal publications, and (2) professional meeting presentations. Service was divided into three subareas: (1) university/college service, (2) department service, and (3) professional service. A 5-point rating scale ranging from "outstanding" (5) to "poor" (1) was used to measure faculty performance on each of the seven appraisal dimensions (i.e., subareas).

After completing the activity report and self-ratings, faculty members returned the appraisal documents to their respective chairpersons.

Each chairperson then rated faculty members in their department using an identical rating form, as described above. They also completed narrative essays, which described their overall impression of each faculty member's performance. In conducting this evaluation, chairpersons not only used the information contained in individual faculty members' self-appraisals, they also used their own observations, as well as information provided by students and other faculty members. [2]

Finally, the chairpersons' ratings, along with their comments were returned to each faculty member. If faculty members disagreed with these ratings or comments, they could discuss this disagreement with their chairpersons. If they still disagreed with the evaluation, they could write a rebuttal and attach it to their performance evaluation forms. These documents were then submitted to the dean of the college for further administrative action.

For example, for student advisement, faculty members were asked to provide the names of the students advised, whether she/he served on the students' committees, each student's current status, and the time period in which the students were advised. It was clear to the faculty members that the contents of their reports could be independently verified on the basis of the above documentation.

[2]The information available to the chairpersons but unavailable to individual faculty members included the activity reports completed by other faculty members. This comparative information might affect chairpersons' evaluation of individual faculty members. Moreover, students who had grievances with faculty members could bring their cases directly to chairpersons. These cases might or might not be reflected in student evaluations of teaching, which were conducted at the end of each semester. In addition to the "formal" information, chairpersons arguably had greater access to "grapevine" information than the average faculty member.

Measures

Four measures were used in this study. The first was self-ratings of performance completed by faculty members. The second was chairperson ratings of performance. As described above, these ratings were based on seven performance dimensions, each with a 5-point measurement scale (1 = poor, 3 = satisfactory, 5 = outstanding). All ratings were obtained from college personnel records.

The third measure consisted of objective performance indicators. These were primarily derived from faculty members' activity reports. They included number of papers under review, chapters published in books, national paper presentations, regional paper presentations, state and local paper presentations, membership on university committees, membership on college committees, journal-reviewing activities, professional activities for academic organizations, honors received, new courses taught, student advisement activities, service on Ph.D. committees, and finally, service on M.S. committees. In addition, chairpersons provided a list of first-, second-, and third-tier journals in their respective fields. This information enabled the researchers to separate the first-tier from second- or third-tier publications. Finally, standardized teaching evaluations collected by the college were used as the criterion measure for teaching performance. Because teaching evaluations were being collected at the same time as the SABPE process, they were unavailable to faculty members and chairpersons and, thus, were independent of both self- and chairperson ratings.

The final measure was a questionnaire survey of user acceptance.[3] To evaluate faculty member reactions to the SABPE process, questionnaires were distributed by the researchers to all faculty members at a point following their annual performance evaluations. The questionnaire first described the SABPE process and then asked faculty members to compare it with a traditional supervisory evaluation (TSE) system in which department chairpersons collected information about faculty member performance and then individually conducted an evaluation. It was noted that a TSE system did not use self-appraisals, and department chairpersons were allowed to make unilateral performance evaluation decisions. Faculty members were then asked to compare the SABPE process with the TSE system on four criteria—fairness, accuracy, comfortableness, and superiority—using a 7-point measurement scale with anchors corresponding to the criteria. For example, the anchors for the fairness criterion ranged from "Definitely less fair than TSE" (1) through "About the same" (4) to "Definitely more fair

[3]The acceptance survey was conducted anonymously. This thus precluded determination of which faculty members were satisfied or dissatisfied with the SABPE system. Since this particular college had used a traditional supervisory evaluation (TSE) system before switching to the SABPE system, most faculty members in our sample had experienced a TSE system.

TABLE 1

Results of Paired t Tests and Means and Standard Deviations of Self- and Chairperson Ratings

Performance dimension	N	Self-ratings M	Self-ratings SD	Chairperson ratings M	Chairperson ratings SD	t Values
Instructional method	77	4.08	.70	4.10	.79	−.31
Instructional support	78	3.92	.79	3.80	.83	1.79
Journal publications	67	3.31	1.38	3.27	1.41	.65
Meeting presentations	64	3.39	1.45	3.44	1.44	−.60
University & college service	57	3.65	.83	3.67	.91	−.21
Department service	73	3.90	.82	3.80	.83	1.38
Professional service	65	3.91	.82	3.79	.89	1.59

Note. Paired *t* tests showed no significant difference in means between self- and chairperson ratings.

than TSE" (7). In addition, faculty members were asked to judge on a 3-point scale (1 = SABPE, 2 = about the same, 3 = TSE) which approach they deemed more effective for evaluating their performance. Seventy faculty members responded to this questionnaire for a response rate of 79%. Furthermore, all department chairpersons were interviewed by the researchers.

The narrative essays completed by department chairpersons were unavailable for this research.

Results

Table 1 presents the means and standard deviations for faculty member self-ratings and chairperson ratings on each of the seven performance dimensions (i.e., activity subareas) described above. The missing data for some performance dimensions (e.g., university and college service, meeting presentations, and journal publications) were mainly due to the fact that full-time instructors in our sample ($n = 15$), whose primary responsibilities were teaching and advising students, were not evaluated on research and service beyond the departmental level. Paired *t* tests were conducted to determine if the ratings were different across the two types of raters. None of the differences were significant. This suggests that self- and chairperson appraisals, as collected through the SABPE process, are equally lenient (i.e., not significantly different).

Chi-square tests were conducted to establish if the standard deviations of the ratings for the seven performance dimensions were significantly different across the two rater types. Again, none of the differences were significant, suggesting that the self- and chairperson ratings were, in general, equally dispersed.

TABLE 2

Multitrait-multimethod Matrix for Faculty Performance Ratings

	Self-ratings							Chairperson ratings					
Ratings	A	B	C	D	E	F	G	A	B	C	D	E	F
Self-Ratings													
B	.52												
C	.14	.26											
D	.08	.18	.74										
E	.50	.70	.48	.31									
F	.51	.73	.39	.28	.72								
G	.45	.36	.32	.38	.54	.47							
Chairperson Ratings													
A	.53	.25	.16	.10	.31	.27	.38						
B	.41	.70	.32	.27	.59	.55	.29	.39					
C	.02	.15	.92	.72	.23	.25	.21	.12	.25				
D	.01	.08	.75	.91	.11	.18	.23	.14	.24	.77			
E	.29	.46	.30	.14	.74	.48	.43	.29	.63	.17	.18		
F	.25	.48	.38	.31	.46	.66	.49	.29	.51	.31	.29	.69	
G	.34	.36	.44	.51	.43	.39	.74	.36	.42	.42	.50	.46	.60

Notes: Convergent validities are underscored. A = instructional method; B = instructional support; C = journal publication; D = meeting presentation; E = university service; F = department service; G = professional service. Sample size varies from 58 to 81 due to missing data.
$r \geq .26$, $p < .05$; $r \geq .34$, $p < .01$.

Table 2 presents the correlations between the self- and chairperson ratings for the seven performance dimensions. Convergent validity coefficients (shown on the square matrix diagonal) for all seven performance dimensions were significant (range = 0.53 to 0.92, $M = 0.78$). The strongest correlations were obtained for journal publications and meeting presentations, r's = 0.92 and 0.91, respectively. The weakest correlations were observed on instructional method and department service, r's = 0.53 and 0.66, respectively. These results indicate that self- and chairperson ratings of faculty member performance collected through the SABPE process largely agreed.

Self- and chairperson ratings of performance were then separately regressed on appropriate objective indicators to assess their criterion-related validity. Table 3(a) and (b) show the results of this analysis for the two scholarship dimensions evaluated. Note that in this table zero-order correlations are reported under column r, and standardized regression coefficients (beta) under column β. While the former indicates the degree of association between an objective indicator and performance ratings, the latter represents a rough estimate of the relative contributions of the objective indicators in predicting self- or supervisor ratings for a given performance dimension. For journal publications, the objective indicators include number of publications in first-tier journals, number of publications in second- or third-tier journals, number of unrefereed publications, number of papers accepted for publication, and number of papers under review. The multiple correlations were .63 for self-ratings and .70 for chairperson ratings (p's $< .01$). Number of first-tier publications and number of papers under review are the most important indicators of scholarly publications. For professional presentations, the indicators include number of national/international presentations, regional presentations, and state/local presentations. The multiple correlations were .59 for self-ratings and .61 for chairperson ratings (p's $< .01$). National/international presentations and regional presentations were the most important indicators of performance in this dimension. Table 3(c) and (d) show the results of multiple regression analyses for instructional method and instructional support. While instructional method focuses on classroom teaching performance, instructional support includes teaching-related activities that are typically performed outside the classroom (e.g., student advisement and serving on graduate committees). For both dimensions the multiple correlations were significant for self- and chairperson ratings. Student teaching evaluations and number of new courses taught were the major performance indicators for instructional method. Number of Ph.D. committees was the most important performance indicator for instructional support.

Table 3(e) and (f) contain results for the service dimensions. For professional service, performance indicators included journal-related activities

TABLE 3

Relationship Between Performance Ratings and Criterion Measures for Performance Dimensions

Performance dimension/criterion measure	Performance rating: Self- β	Self- r	Chairperson β	Chairperson r
(a) Journal publication				
Refereed publications				
(1st tier)	.33**	(.45**)	.34**	(.54**)
(2nd or 3rd tier)	.04	(.41**)	.12	(.51**)
Nonrefereed publications	−.05	(.11)	.06	(.24**)
Papers accepted for publication	.17	(.42**)	.15	(.43**)
Papers under review	29*	(.50**)	.29*	(.55**)
Chapters published	.12	(.17)	.06	(.12)
Books published	−.02	(.10)	−.01	(.16)
R		.63**		.70**
(b) Professional presentations				
National/international	.43**	(.50**)	.43**	(.50**)
Regional	.28**	(.34**)	.33**	(.38**)
State/local	.14	(.25*)	.15	(.26*)
R		.59**		.61**
(c) Instructional method				
Student evaluations	.21	(.19)	.45**	(.45**)
New courses taught	.34**	(.30*)	.22*	(.14)
R		.37**		.47**
(d) Instructional support				
Student organizations participated	.25	(.31**)	.18	(.28*)
Ph.D. committees	.31**	(.47**)	.44**	(.46**)
M.S. committees	.20	(.33**)	−.04	(.15)
R		.54**		.50**
(e) Professional service				
Journal related activities	.21	(.19)	.34**	(.31**)
Meeting activities	.12	(.10)	.09	(.08)
Elected positions	.27*	(.31*)	.30*	(.35**)
Honors received	.19	(.27*)	.20	(.30*)
R		.43**		.52**
(f) College/university service				
College committees	.37**	(.43**)	.45**	(.50**)
University committees	.38**	(.44**)	.35**	(.41**)
R		.57**		.61**

Note: Sample size varies from 58 to 81 in the analyses due to missing values; r indicates zero-order correlations; β indicates standardized regression coefficients; R indicates multiple correlations.

*$p<.05$; **$p<.01$

(i.e., serving as ad hoc reviewers or on journal editorial boards), meeting activities (i.e., serving as discussants, session chairs, or reviewers for professional meetings), elected positions, and honors received. Among these indicators, journal-related activities and elected positions were important indicators of performance for professional service. Committee memberships were the major means by which faculty could provide college and university service. As expected, number of committee memberships at either level was a significant indicator of performance for college/university service. The multiple correlations for these two dimensions were significant for both self- and chairperson ratings.

Each of the preceding dimensions had a set of clear and unique performance indicators. This, however, was not true for department service. This dimension was not well defined by the appraisal instrument, and the activity report contained no performance indicators uniquely pertaining to it. Consequently, this dimension was excluded from the ensuing validity analysis.

Table 4 presents faculty member reactions to the SABPE process as compared with a TSE system. Results indicated that 70% or more of the faculty surveyed believed the SABPE process was fairer, more accurate, and superior to a TSE system. Additionally, some 60% of the faculty indicated greater comfort with the SABPE process than with a TSE system. When faculty members were asked to compare the relative effectiveness of the SABPE and TSE systems, 75.4% reported SABPE was more effective, 18.4% reported they were indifferent, and 6.2% reported TSE was more effective. These results indicated that a majority of the faculty responded very favorably to the SABPE process and believed it to be superior to a TSE system for evaluating performance.

Faculty acceptance of the SABPE process was also reflected by the low number of grievances filed by individual faculty members. Only one faculty member appealed a chairperson's evaluation.

Interviews with six department chairpersons showed that four favored the SABPE process over a TSE system. Two thought that both were equally satisfactory. Of these two, one was an acting department chairperson and felt unable to make a definitive comparison between the two systems. The other regularly used MBO for performance evaluation purposes and indicated that a SABPE process did not have any impact on this practice. The department chairpersons noted that from their perspective the SABPE process had three distinct advantages over a TSE system: (1) they felt less defensive in the performance interviews; (2) they felt more confident in their evaluations and, thus, more comfortable in using them to justify their administrative decisions; and (3) they believed that discrepancies between self- and chairperson ratings could be used to clarify disagreements concerning performance standards.

TABLE 4

Results of Faculty Responses Comparing the TSE System with the SABPE Process

Criterion	TSE preferred	About the same	SABPE preferred	Mean	*SD*
Fairness	5.7%	18.6%	75.7%	5.49	1.38
Comfortable	17.1%	21.4%	61.5%	5.11	1.59
Accuracy	7.1%	22.9%	70.0%	5.20	1.37
Superior	8.5%	14.3%	77.2%	5.41	1.35
Overall evaluation	6.2%	18.4%	75.4%	1.31	0.58

Note: TSE = Traditional Supervisory Evaluation System; SABPE = Self-appraisal Based Performance Evaluation Process.

$N = 70$.

Discussion

This study found that in a self-appraisal-based performance evaluation (SABPE) process, self-ratings were highly congruent with supervisor ratings. Self-ratings were found to be just as dispersed and no more lenient than supervisor ratings on the seven performance dimensions. In addition, significant convergent validity coefficients were obtained between self- and supervisor ratings for seven performance dimensions.

The generally high congruency between self- and supervisor ratings may be due to their lack of independence. That is, supervisor ratings were obtained after the chairpersons had reviewed self-ratings. Chairpersons may have agreed with their subordinates in order to avoid potential confrontations. This may be especially true when self-ratings were judged too lenient, but conclusive counter evidence did not exist. This, however, does not imply that supervisors would go along with self-ratings when they perceived such ratings to be unduly lenient (or harsh) in view of the evidence contained in individual faculty member activity reports.

Indeed, there are other factors in a SABPE process that may have contributed to the high congruency between self- and supervisor ratings. First, in the SABPE process self- and chairperson ratings were largely made on the basis of a common pool of performance information generated in the self-documentation process (i.e., activity reports). Second, the performance dimensions used in this study focused on specific outcomes or activities as opposed to broadly defined traits. Thus, they were less subject to rating biases. Previous research shows that self-appraisals on ambiguous attributes are more susceptible to leniency bias than are more concrete attributes (e.g., Felson, 1981). Third, as noted earlier, there is evidence that self-ratings tend to be more accurate when self-raters expect that their ratings will be validated against independent criterion measures (Bauman & Dent, 1982; Farh & Werbel, 1986; Mabe & West, 1982). In the SABPE process, this expectation is patently evident. Not only do faculty members expect that

their self-ratings will be reviewed by their chairpersons, but they are called upon to provide evidence to justify their self-ratings.

It was also found that self- and chairperson ratings were significantly correlated with appropriate criterion measures. Multiple correlations ranged from .37 to .63 for self- and .47 to .70 for chairperson ratings. These results are likely to be conservative estimates of validity because the objective indicators used in this study were mostly quantitative in nature, with qualitative performance aspects being underrepresented. For example, for journal publications, no allowance was made for single versus multiple author publications, nor was article length considered. These factors were probably taken into account in the rating process by our raters but were not explicitly included in the study's criterion measures. Nevertheless, our results indicate that performance ratings obtained in a SABPE process are fairly accurate.

Moreover, user acceptance of the SABPE process was very high. Supervisors believed that, among other benefits, the SABPE process reduced defensiveness in performance interviews and helped resolve disagreements between subordinates and supervisors. These results are consistent with the findings reported in the earlier cited GE experiment (Bassett & Meyer, 1968). In addition, subordinates viewed the SABPE process as more accurate, fair, comfortable, and in general, more effective than the traditional supervisory evaluation. This is not surprising because the SABPE process allows subordinates the opportunity (1) to provide information to influence supervisory evaluations, (2) to engage in two-way communication during performance interviews, and (3) to rebut unfavorable evaluations. These characteristics are central to the concept of process control in the theory of procedural justice (Thibaut & Walker, 1975) and have been found to be key determinants of perceived fairness in performance evaluation (Greenberg, 1986).

In summary, our findings indicate that when self-appraisals are used in the SABPE process, they are highly correlated with supervisor ratings and moderately correlated with objective performance indicators. Unlike peer evaluation procedures, where user acceptance is reported low (e.g., Cederblom & Lounsbury, 1980; Love, 1981), most SABPE users clearly indicate that they prefer the SABPE process to traditional supervisor-prepared performance evaluations. Judging from the above results, we conclude that, at least for the focal sample, the SABPE process has proven to be a successful alternative to traditional supervisor-prepared performance evaluations.

To what extent are the results from this study applicable to other administrative settings? The answer to this question, of course, depends on the nature of the settings to which these results are to be generalized. We believe that the results from this study are likely generalizable to other higher-education institutions, especially those placing a strong emphasis on

research. By contrast, when the SABPE process is applied to traditional work settings, three situational factors should be considered. The first factor to consider is management style. As described earlier, the SABPE process is essentially a method that invites greater employee involvement in performance appraisals. Like similar techniques, it is best applied in organizations having a participative management style. Therefore, it is doubtful that an organization should try the SABPE process unless it has initiated or is ready to initiate a democratic or participative style. The second factor is the extent to which employees work independently. The SABPE process is ideally suited to situations in which employees frequently engage in independent work under minimal supervision. Under such circumstances, self-appraisals are likely to contain unique information and, thus, to be viewed as valuable additions to traditional supervisor appraisals. A final factor to consider is the clear definition of performance criteria. To facilitate interrater agreement, an organization should not only try to define performance criteria unequivocally, but it should also specify results or activities that constitute success or failure for each performance dimension. This will help employees prepare valid activity reports on which both self- and supervisor appraisals are to be based.

Future research is needed to determine if the SABPE process is a viable complement or, perhaps, alternative to supervisor-prepared evaluations in traditional work settings. Moreover, the SABPE process as described in this study is only one approach to self-appraisal-based performance evaluation. Future researchers should explore other approaches to and applications of the SABPE process. As an example of another application, the SAPBE process could be applied to evaluating employee development needs. For instance, supervisors could be asked to independently appraise subordinate training needs before they have access to self-appraisals. Only with more research will we know the full benefits and limitations of using self-appraisals as a tool in the performance appraisal process.

REFERENCES

Anderson CD, Warner JL, Spencer CC. (1984). Inflation bias in self-assessment examinations: Implications for valid employee selection. *Journal of Applied Psychology, 69*, 574–580.

Bassett GA, Meyer HH. (1968). Performance appraisal based on self-review. PERSONNEL PSYCHOLOGY, *21*, 421–430.

Bauman KE, Dent CW. (1982). Influence of an objective measure on self-reports of behavior. *Journal of Applied Psychology, 67*, 623–628.

Borman WC. (1974). The rating of individuals in organizations: An alternate approach. *Organizational Behavior and Human Performance, 12*, 105–124.

Carroll SJ, Schneier CE. (1982). *Performance appraisal and review systems: The identification, measurement, and development of performance in organizations.* Glenview, IL: Scott, Foresman.

Cederblom D, Lounsbury JW. (1980). An investigation of user acceptance of peer evaluations. PERSONNEL PSYCHOLOGY, *33*, 567–579.

DeNisi AS, Shaw JB. (1977). Investigation of the uses of self-reports of abilities. *Journal of Applied Psychology*, *62*, 641–644.

Downs S, Farr RM, Colbeck L. (1978). Self-appraisal: A convergence of selection and guidance. *Journal of Occupational Psychology*, *51*, 271–278.

Farh JL, Werbel JD. (1986). The effects of purpose of the appraisal and expectation of validation on self-appraisals leniency. *Journal of Applied Psychology*, *71*, 527–529.

Felson RB. (1981). Ambiguity and bias in the self-concept. *Social Psychology Quarterly*, *44*, 64–69.

Fletcher C. (1986). The effects of performance review in appraisal: Evidence and implications. *Journal of Management Development*, *5*(3), 3–12.

Greenberg J. (1986). Determinants of perceived fairness of performance evaluations. *Journal of Applied Psychology*, *71*, 340–342.

Heneman HG III. (1974). Comparisons of self and superior ratings of managerial performance. *Journal of Applied Psychology*, *58*, 638–642.

Klimoski RJ, Hayes NJ. (1980). Leader behavior and subordinate motivations. PERSONNEL PSYCHOLOGY, *33*, 543–555.

Landy FJ, Farr JL. (1980). Performance rating. *Psychological Bulletin*, *87*, 72–107.

Latham GP, Wexley KN. (1981). *Increasing productivity through performance appraisal.* Reading, MA: Addison-Wesley.

Levine EL, Flory A III, Ash RA. (1977). Self-assessment in personnel selection. *Journal of Applied Psychology*, *62*, 428–435.

Love KG. (1981). Comparison of peer assessment methods: Reliability, validity, friendship bias, and user reaction. *Journal of Applied Psychology*, *66*, 451–457.

Mabe PA II, West SG. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, *69*, 280–296.

Meyer HH. (1980). Self-appraisal of job performance. PERSONNEL PSYCHOLOGY, *33*, 291–295.

Shrauger JS, Osberg TM. (1981). The relative accuracy of self-predictions and judgments by others in psychological assessment. *Psychological Bulletin*, *90*, 322–351.

Thibaut J, Walker L. (1975). *Procedural justice: A psychological analysis.* Hillsdale, NJ: Erlbaum.

Thornton GC. (1980). Psychometric properties of self-appraisals of job performance. PERSONNEL PSYCHOLOGY, *33*, 263–271.