

Performance Appraisal: Some Unintended Consequences

by
Mark A. McBriarty

Few areas of personnel management have received more attention than performance appraisal, and few have remained as shrouded in controversy and contradiction. After ingesting even a small portion of the literature on the subject, one typically ends up with an acute case of mental indigestion. Faced with a plethora of opinions, managers today are forced to wrestle with often-conflicting prescriptions for the "one best way" to devise a useful and effective way to appraise employee performance. The diversity evidenced in both the form and substance of their attempts suggests the extent of the disagreements involved.¹ What is perhaps most disconcerting is that, despite all of the cautions about the complexity of the process, managers frequently succumb to the lure of simplicity and convenience that still beckons us, like the legendary Lorelei, to the shoals of organizational disaster.² Too often, they forget the admonitions Wallace Sayre offered four decades ago and yield to such temptation, falling victim to the "...canonization of rudimentary techniques, fascinated more by procedural rule and techniques rather than purpose or results...steadily more dependent upon a cold, impersonal, rigid quantification of human ability and worth."³ The unfortunate result is that many managers still find themselves flying by the seat of their pants, without any real appreciation of what their performance appraisal system is doing for, or to, their organization.⁴

This article makes no attempt to carve a path through the performance appraisal jungle, leading to the "one best way." That course, if it exists, remains to be charted. Rather, the purpose here is to describe some observations of the unintended consequences of one approach to performance appraisal attempted by the Department of

Dr. McBriarty is currently an Assistant Professor of Public Administration at the University of Texas at San Antonio, specializing in human resources management. He holds the MBA and DPA degrees from George Washington University and has had over 15 years experience as a public practitioner working with the design and implementation of federal management improvement initiatives. While with the Department of the Air Force, in Washington D.C., he managed the Air Force Cost Reduction Program and served as the Director of the Air Force Management Review and Improvement Program. More recently, he held the position of Chief of Management Improvement Programs at the Air Force Manpower and Personnel Center in San Antonio, Texas. he has conducted extensive field research in each of the major agencies and departments in Washington, D.C. concerning their reactions to mandated management improvement initiatives and is the author of several papers addressing problems in public sector human resource management.

the Air Force from 1974 to 1978.⁵ Hopefully, the observations offered will add to the inventory of understanding about the organizational dynamics that need to be given careful consideration in the design and implementation of any performance appraisal system.

THE AIR FORCE APPRAISAL SYSTEM





Formal appraisal of the performance and potential of individual officers has been a central component of personnel management in the Department of the Air Force since it was chartered as a separate service in 1947. Understandably, the appraisal system has been the focus of continual concern since the annual performance reports it generates serves as the principal basis for decisions concerning promotions, job assignments, training, and retention in the service. In recent years, these reports have assumed added significance as decisions in these areas have been increasingly centralized in boards of officers convened at the Military Personnel Center—officers who typically have little or no personal knowledge of the performance or abilities of the individuals affected by their decisions. In such a situation, where an unfavorable board action, based on a collection of lifeless performance reports, can have catastrophic career implications, the potential impacts of the appraisal system on the human organization are significant.

While there have been a number of modifications to the officer evaluation system over the years, most of these have involved relatively minor changes in reporting formats and administrative procedures, leaving the substance of the process basically unchanged.⁶ Essentially, the system has remained a hybrid of conventional techniques and approaches described in the literature and commonly found in public sector organizations.⁷ It requires an annual written performance report on each officer below the rank of general, prepared by the immediate supervisor and endorsed by his or her superior. The reporting format and procedures are centrally designed by the Military Personnel Center in San Antonio, Texas, for standardized application throughout the Air Force. Appraisal requirements apply indiscriminately to all officers, regardless of occupational specialty, level of responsibility, or locus in the organizational hierarchy. The same appraisal instrument is used to evaluate the performance of pilots, doctors, chaplains, and other staff officers, whether assigned to a headquarters or field activity, regardless of rank or longevity.

THE APPRAISAL INSTRUMENT

Although there have been several editions of the annual officer appraisal instrument, they have consistently been designed to generate both descriptive and comparative data concerning personal traits and job performance.⁸ The two versions of the form (AF Form 77) used from 1966 to 1974 had only minor editorial differences and exemplify this hybrid approach. They consisted of nine sections designed to collect data for a number of different purposes. Sections I and II captured ratee identification data including name, grade, organization, functional specialty, job title, and

FIGURE 1

III. RATING FACTORS (Consider how this officer is performing on his job.)					
1. JOB CAPABILITY					
<p>NOT</p>  <p>OBSERVED</p>	<input type="checkbox"/> HAS GAPS IN FUNDAMENTAL KNOWLEDGE AND SKILLS OF HIS JOB.	<input type="checkbox"/> HAS A SATISFACTORY KNOWLEDGE AND SKILL FOR THE ROUTINE PHASES OF HIS JOB.	<input type="checkbox"/> HAS EXCELLENT KNOWLEDGE AND IS WELL SKILLED ON ALL PHASES OF HIS JOB.	<input type="checkbox"/> HAS AN EXCEPTIONAL UNDERSTANDING AND SKILL ON ALL PHASES OF HIS JOB.	<input checked="" type="checkbox"/> HAS A FAR-REACHING GRASP OF HIS ENTIRE BROAD JOB AREA. AUTHORITY IN HIS FIELD.
2. PLANNING ABILITY					
<p>NOT</p>  <p>OBSERVED</p>	<input type="checkbox"/> RELIES ON OTHERS TO BRING PROBLEMS TO HIS ATTENTION. OFTEN FAILS TO SEE AHEAD.	<input type="checkbox"/> PLANS AHEAD JUST ENOUGH TO GET BY IN HIS PRESENT JOB.	<input type="checkbox"/> IS A CAREFUL EFFECTIVE PLANNER. ANTICIPATES AND TAKES ACTION TO SOLVE PROBLEMS.	<input type="checkbox"/> CAPABLE OF PLANNING BEYOND REQUIREMENTS OF THE PRESENT JOB: SEES THE BIG PICTURE.	<input checked="" type="checkbox"/> CAPABLE OF TOP LEVEL PLANNING. A HIGH CALIBER THINKER AND PLANNER.
3. EXECUTIVE MANAGEMENT					
<p>NOT</p>  <p>OBSERVED</p>	<input type="checkbox"/> IS A POOR ORGANIZER. DOES NOT REALLY MAKE EFFECTIVE USE OF MATERIAL OR MANPOWER.	<input type="checkbox"/> MAINTAINS ORDINARY EFFICIENCY OF OPERATION. CONTROL COULD BE IMPROVED.	<input type="checkbox"/> GIVES ECONOMY OF OPERATION CAREFUL ATTENTION. MAKES WISE USE OF MANPOWER AND MATERIAL.	<input type="checkbox"/> MAINTAINS EFFECTIVE ECONOMY. CAREFULLY WEIGHS COST AGAINST EXPECTED RESULTS.	<input checked="" type="checkbox"/> HIGHLY SKILLED IN BALANCING COST AGAINST OPTIMUM EFFECTIVENESS.
4. LEADERSHIP					
<p>NOT</p>  <p>OBSERVED</p>	<input type="checkbox"/> OFTEN WEAK IN COMMAND SITUATIONS. AT TIMES UNABLE TO EXERT CONTROL.	<input type="checkbox"/> NORMALLY DEVELOPS FAIRLY ADEQUATE CONTROL AND TEAMWORK.	<input type="checkbox"/> CONSISTENTLY A GOOD LEADER. COMMANDS RESPECT OF HIS SUBORDINATES.	<input type="checkbox"/> EXCEPTIONAL SKILL IN DIRECTING OTHERS TO GREAT EFFORT.	<input checked="" type="checkbox"/> LEADERSHIP QUALITIES REFLECT POTENTIAL FOR HIGHEST LEVEL.

a brief, generalized job description covering key duties and responsibilities. Section III was devoted to an assessment of performance in terms of eight rating factors including personal traits such as leadership, adaptability, and judgment, along with several job related items such as job knowledge, planning, resource management, written and oral communication. In each of these areas, appraisers were required to mark a block on a scale from one (lowest), to five (highest), indicating how well the individual's performance approximated the brief descriptions accompanying each block. (See Figure 1). Section IV asked for an evaluation of the individual's military bearing and behavior using the same type of scale. Section V was an overall evaluation of the individual compared to his peers, rated on a ten point scale ranging from unsatisfactory to "absolutely superior", with specific narrative justification required to ratings below average or outstanding/superior. Section VI was an evaluation of promotion potential, that suggested the timing of promotion, relative to peers. (See Figure 2). Section VII of the report provided space for the rater's narration of the critical facts and achievements relating to the individual's performance during the period covered by the evaluation. Section VII included rater identification data, and

FIGURE 2

VI. PROMOTION POTENTIAL	
1. DOES NOT DEMONSTRATE A CAPABILITY FOR PROMOTION AT THIS TIME. <input type="checkbox"/>	2. PERFORMING WELL IN PRESENT GRADE. SHOULD BE CONSIDERED FOR PROMOTION ALONG WITH CONTEMPORARIES. <input type="checkbox"/>
3. DEMONSTRATES CAPABILITY FOR INCREASED RESPONSIBILITY. CONSIDER FOR ADVANCEMENT AHEAD OF CONTEMPORARIES. <input type="checkbox"/>	4. OUTSTANDING GROWTH POTENTIAL BASED ON DEMONSTRATED PERFORMANCE. PROMOTE WELL AHEAD OF CONTEMPORARIES. <input checked="" type="checkbox"/>

Section IX solicited comments by the Endorsing Official, typically the rater's immediate supervisor.

THE CONTROLLED APPRAISAL

The Impetus For Change

As indicated earlier, the appraisal instrument described briefly above was the centerpiece of the Air Force officer evaluation system for the better part of a decade. While it is difficult to scientifically assess the utility of the report in terms of appraisal objectives commonly identified in the literature, a cursory review of the components of the instrument suggests some inherent limitations.⁹ For one thing, the job descriptions entered on the evaluation report were usually generalized and abbreviated to a point where they said little about the specific organizational expectations associated with the job performed by the individual. Hence they were of questionable value in communicating organizational objectives to the incumbent or in clarifying the relationship of his particular performance to the efforts of the rest of the organization. Similarly, the subjectivity and imprecision of the rating factors used made them of questionable reliability, offering little in the way of quantifiable benchmarks against which performance could be objectively measured. Consequently, identification of specific areas of performance in need of improvement was difficult at best. In addition, some of the rating factors used had only minor relevancy to the job many individuals were required to perform. For example, assessing the leadership abilities of an individual with no subordinates is, at face value, a questionable practice. Equally as troublesome was the requirement to render a composite, scalar evaluation of each ratee, comparing him as a whole to someone else. As Stahl suggests, "This is a most doubtful if not actually dangerous procedure, since an individual...usually finds invidious and ego shattering any attempt to pin a label on his total worth."¹⁰

A similar observation would seem in order concerning the aggregate assessment of promotion potential required in Section VI of the report. As a number of authors

have agreed, the combination of this kind of prospective judgement with evaluations of past performance tends to dilute the effectiveness of the entire assessment effort.¹¹ Lastly, the heavy reliance on the undocumented narrative of specific achievements included in the report, offered an open opportunity for some fairly creative writing that usually reflects the talent and energy of the author as much as the real accomplishments of the ratee.

Taken together, the shortcomings mentioned above might well have been ample cause for concern when the system was evaluated in the early 1970s.¹² However, the focus of efforts to revise the system at that time centered on quite a different problem; the apparent inability of the appraisal system to provide management a meaningful differentiation among the members of the officer corps in terms of their demonstrated abilities and future potential. By 1973, over 90% of the officer corps was receiving top block ratings in both performance and promotion potential. Whether this trend was more the result of a general inclination toward leniency to avoid conflict and the distastefulness of "playing God", or the reflection of the traditional values of teamwork, conformity, and esprit, peculiar to the military, remains argumentative.¹³ Regardless of the real reason behind this inflationary pattern, it was decided, based largely on anecdotal evidence, that promotion and other central selection boards were somehow being deprived of critical information; the inference being that the wrong people were being promoted and otherwise recognized. While this contention was never really demonstrated, it became the dominant influence behind a drive to stamp out inflation of ratings and force the level of differentiation that would satisfy the information needs of the personnel management system.¹⁴ Procedural standardization, administrative simplicity, and technical sufficiency became the by-words of reform, and the forced distribution of ratings emerged as the best way to eliminate the systematic biases that had crept into the system.¹⁵

The Forced Distribution

The changes implemented in 1974 did little to alter the basic approach to officer appraisal, and essentially retained the structure of the evaluation instrument described earlier. With minor semantic changes and some modifications to the rating scale, the "Performance Factors" section of the report remained largely the same, as did the sections covering job description and ratee identification data. The only drastic change concerned Section V of the new report which required an evaluation of the overall potential of the ratee, in terms of his capability, relative to his peers, to assume increased responsibilities. Raters and endorsing officials were required to record their assessments of potential on a scale ranging from one (highest) to six (lowest) by placing an "X" in the appropriate block on the scale. (See Figure 3). On the surface, this change might appear to be mainly cosmetic, a fine tuning of the evaluation instrument. In fact, however, the guidance that accompanied the revision made it perhaps the single most significant step in the evolution of the appraisal system.

The logic behind the new rules concerning the forced distribution of ratings was

FIGURE 3

V. EVALUATION OF POTENTIAL

TOP BLOCK CONTROLLED

Evaluate the ratee's capability, relative to that of officers in the same grade in the group being evaluated, for expanded/more diverse responsibility. Indicate your rating by placing an "X" in the designated portion of the appropriate block.

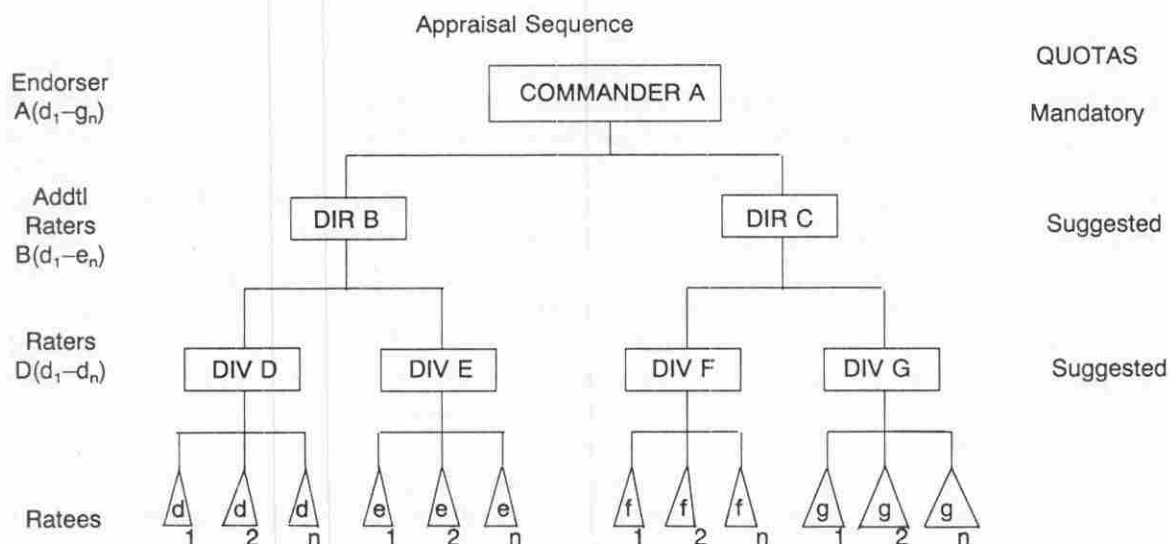
<table border="1" style="width: 100%; height: 20px;"> <tr> <td style="width: 33%;"></td> <td style="width: 33%;"></td> <td style="width: 33%;"></td> </tr> </table> <p style="text-align: center; font-size: small;">RATER ADDN REVR RATER</p>				<table border="1" style="width: 100%; height: 20px;"> <tr> <td style="width: 33%;"></td> <td style="width: 33%;"></td> <td style="width: 33%;"></td> </tr> </table> <p style="text-align: center; font-size: small;">RATER ADDN REVR RATER</p>				<table border="1" style="width: 100%; height: 20px;"> <tr> <td style="width: 33%;"></td> <td style="width: 33%;"></td> <td style="width: 33%;"></td> </tr> </table> <p style="text-align: center; font-size: small;">RATER ADDN REVR RATER</p>				<table border="1" style="width: 100%; height: 20px;"> <tr> <td style="width: 33%;"></td> <td style="width: 33%;"></td> <td style="width: 33%;"></td> </tr> </table> <p style="text-align: center; font-size: small;">RATER ADDN REVR RATER</p>			

← Lowest

↑ Highest

intuitively simple. It assumed that performance potential, like everything else in life, varies along some continuum that should have some distributional properties to it.¹⁶ By requiring ratings of potential to follow a distributional pattern, the system would insure that the "high" and "low" performers would be identified. While such reasoning seems compelling, it afforded scant recognition to the caution offered by L. Cummings and others about the potential errors inherent in such an approach.¹⁷ The distribution in this case was established by the staff at the Air Force Military Personnel Center and limited ratings of potential to 22% in the highest block, 28% in the next highest, with the remaining 50% to be distributed over the remaining blocks. Within this distribution, percentage quotas were strongly encouraged for raters (first line supervisors) and additional raters (second line supervisors). For the endorsing official, usually the final reviewing authority, the distribution quotas were mandatory. Figure 4 is a simplified illustration of the appraisal sequence in a standard organization. In this example, the rating officials, Division Chiefs (D,E,F,G) prepare the appraisal reports for all of their subordinate officers in a particular grade ($d_1 - g_n$) during the same month of the year. They were strongly urged, although few did, to make their ratings of potential conform to the prescribed distribution. Reports were then passed to the next echelon, the Directorate level, where the additional raters (B & C) added their comments to the reports along with their assessment of potential. Adherence to the distribution quotas was also recommended at this level. Reports then flowed to the Endorsing Official (A) who was required to add his evaluation of potential, limiting his ratings to the quotas in the prescribed distribution, i.e. 22% in the top block, 28% in the next, and 50% among the rest, (3 and below). The sequence was repeated for each of the officer grades at different times throughout the year. Rater, Additional Rater, and Endorser responsibilities would be shifted up or down in the hierarchy, depending upon the position of the officer being evaluated. For example, in the illustration, when Division Chiefs (D & G) were to be evaluated,

FIGURE 4



Directors (B & C) would serve as raters, Commander (A) would be the Additional Rater, and his supervisor would serve as the Endorsing Official. In all instances, every officer received at least one "controlled" evaluation of potential each year.

There were a number of important aspects of the forced distribution requirement that are not evident from this simple illustration. For one thing, the sheer numbers involved presented a formidable problem for endorser. They were faced with the task of making relative evaluations of dozens of officers scattered throughout the organization by registering their aggregate assessments with a single "X" on the evaluation form. As O.G. Stahl points out, the assumption that supervisors can somehow "...analyze all of the faults and strengths of an employee, add them up in some fashion, and come to some neat overall conclusion...expressed by a percentage or numerical value...that makes it possible to compare him as a whole to someone else..." is a dangerous procedure.¹⁸ His counsel seems particularly germane in this situation, given the inverse relationship that usually exists between the Endorser's first hand knowledge of ratee's abilities and the number of individuals being evaluated.

Compounding the problem of numbers was the fact that Endorsers were frequently separated from the individuals being appraised by organizational or geographical distances, having little or no opportunity to actually observe performance. It was not uncommon for an individual to have his appraisal form endorsed by someone he had never met or even talked to. For example, the vagaries of the chain of command might dictate that a weather officer, serving in the south Pacific, would have his annual appraisal report endorsed by a superior who was stationed in Illinois. To make matters were difficult, it made no difference that the endorser in Illinois may have

been in his position for only a matter of days before being required to endorse performance reports on subordinate officers in the field. Yet the endorsement proved to be the most critical component of the entire appraisal process.

In addition, the use of distribution quotas by endorsing officials was universally prescribed for all reports, regardless of the organizational placement of the individuals being evaluated. Appraisals of individuals hand picked for the most important and difficult jobs at Department Headquarters were subject to the same restrictions as those rendered on officers assigned to lower priority units, despite the fact that by design, quality was concentrated in these selectively manned units.¹⁹ What this meant was that 50% of the officers in these units had to be rated average or below when compared to their peers in the rest of the Air Force by their Endorsing Official. Under the rules, half of the Air Force astronauts selected to fly in space would be rated in the lower half of the officer corps as far as their potential was concerned. Magnifying the potential effects of these difficulties, as appraisal results were put into the computer and circulated in the personnel management system, it became clear that the controlled rating of potential was the single most important factor in promotion, assignment, and training decisions. In the pattern suggested by E. Yager, the rest of the appraisal report seemed transparent, regardless of the actual performance it described.²⁰

SOME UNINTENDED CONSEQUENCES

The initial response to these changes was understandably muted and buried in bureaucratic busyness. However, it wasn't long before surveys and interviews with individual officers began to detect an almost universally negative reaction to the forced distribution system; a reaction that threatened serious dysfunctional behavior by a substantial segment of the officer corps. While it is difficult if not impossible to quantify the behavioral impacts of the changes on individual performance, or aggregate organizational effectiveness, some general tendencies were apparent in several critical areas.

Motivation

Recognizing the gaps in our conventional wisdom concerning the relationship between motivation and performance, one would expect little argument with the proposition that constructively motivated managers are a key ingredient in the recipe for organizational success. One would also likely find general agreement with the assertion that, whenever possible, management policies and systems should be supportive of such motivation. The appraisal system described above clearly worked in the opposite direction, focusing on inordinate amount of attention on the basic human concerns about survival, security, and ego maintenance at the expense of the higher order "motivators" of more productive organizational behavior.²¹ With a simple administrative fiat, the comfortable perceptions of self worth and competence, com-

monly enjoyed by the officer corps, and reinforced by years of "outstanding" ratings, were summarily exploded for half of its members, leaving them to cope with the painful realization that the organization now considered their net worth "average or below". The typical response to this experience was predictably negative, and can best be characterized as a combination of pouting, cynicism, alienation, and half-hearted performance similar to that described by Dalton and others.²² Faced with considerable uncertainty about their futures, in terms of promotion and tenure, it was not surprising to see these individuals drift toward the kind of organizational disinvestment noted by Levine & Wolohojian, reviewing the options for alternative employment and doing just enough to get by in the positions they occupied.²³ While it is impossible to determine how many may have eventually left the organization because of their dissatisfaction, there were clear indications of a general reassessment of the balance between the costs and benefits involved with continued association with the organization.²⁴ In the process, motivation for sustained effort and performance was an early casualty.

SUPERVISORY RELATIONS

In addition to these adverse motivational impacts, the controlled distribution system also tended to erode the quality of subordinate/supervisor relationships and undermine the capability of the latter to function effectively.²⁵ In many ways, it intensified the inevitable clash between the subjective words of the supervisor and subordinate that Nalbandian sees as inherent in any appraisal process that officially tells an individual he is not as good as someone else.²⁶ For those individuals told by the "system" that they were somehow inferior to 50% of their peers, such conflict was real and placed considerable strain on their loyalty to and support of their supervisors. They perceived a situation where they were faced with basically two alternatives; either accept the appraisal as a valid assessment of their capabilities, or question the judgement and fairness of their superiors. Since acceptance of relative inferiority is the exception rather than the rule, it was not surprising to find individuals opting for the more comfortable defense found in concluding that their supervisors didn't really know what they were doing. Such opinions were undoubtedly reinforced in those frequent instances where raters and reviewers had only limited contact with the individual being evaluated.²⁷

The controlled distribution system exerted a negative influence on supervisor/subordinate relations in one other important respect. Appraisals of supervisors' performance were subject to the same discipline as those rendered on their subordinates. While results of appraisals were not officially publicized, they quickly became common knowledge through the organizational "grapevine" of informal communication. Subordinates soon realized that their supervisors were, at least 50% of them, judged to be average or below. This realization did little to enhance the image of the supervisor in the eyes of the subordinate and prompts speculation about the degree to which the chain of command was weakened as a result.

Decision Making

Another, and perhaps more subtle, impact of the forced distribution of ratings could be observed in the organizational locus of decision making. Under this appraisal system, it became obvious that one's relative advantage in the competition for limited top block ratings was considerably enhanced by increased "facetime" with the supervisor and reviewing official. Consequently, there was a distinct tendency to inflate the importance or urgency of particular issues in hopes of escalating them upward in the chain of command to a level where visibility would be assured. In one organization, whenever a high ranking official would visit, it was common practice to schedule a rapid sequence of decision briefings for the dignitary so that the maximum number of ratees could get some exposure to the "big boss", whether or not the information they presented was significant enough to warrant such an audience. While such a procedure may have helped some individuals in the competition for top block ratings, it often represented a questionable use of top management's time by forcing high level attention to decisions that could have been made at much lower levels in the organization.

This "visibility syndrome" had another dimension evidenced in the way credit was distributed for what was actually accomplished within an organization. In many instances, rather than letting the individuals who actually did the work get the visibility for their achievements, supervisors were inclined to put their own label on the significant efforts of their subordinates, taking personal credit for the successes and distributing blame for the failures. When a subordinate was permitted to present the results of his efforts upward in the organization, protocol dictated that an "adult supervisor" be present so that visibility could be shared. This practice is not uncommon in large bureaucracies, but it seemed to be encouraged by the competitive nature of the appraisal system.

Competition under the controlled rating system also worked to influence the time horizons of decisions concerning resource utilization and the direction of organizational effort. Under pressure to show dramatic, eye-catching achievements before the next appraisal cycle, supervisors tended to focus their attention and resources on those projects with high potential payoff in the short term. Support for the routine, low profile activities associated with long term organizational growth was less than enthusiastic; something to be avoided by anyone with an eye on the fast track. While other bureaucratic imperatives undoubtedly contributed to this short term orientation, the controlled appraisal system all but demanded adherence to such priorities, whether or not they were in the best interest of the organization over the long term.

Teamwork

By definition, the requirement for a forced distribution of "potential" ratings was a zero sum system in that any change in the relative ranking of individuals had to add to "O". For one individual to win (get a higher rating) a peer had to lose (get a lower rating). This feature made the interaction among peers in any particular or-

ganization intensely personal and competitive. Given the high stakes involved in terms of tenure and promotion, this competition understandably assumed proportions that did little to foster the kind of teamwork and cooperative effort essential to the health of any organization. Rather, it had a decidedly destructive effect, putting individuals in a situation where any assistance to a peer might ultimately prove to be a distinct disadvantage in the next evaluation cycle. Essentially, it had the effect of depressing any motivation to actively support the efforts of one's peers and encourage a mentality where looking out for "number one" was the most rational behavior pattern. The net result was a kind of "dog-eat-dog" combat similar to that noted by F. Thayer in his discussion of response to the competitive rating features of the Civil Service Reform Act.²⁸ In the longer term, such competition would seem to have had an additional negative impact on organizational performance by undermining the ethos of mutual dependency among action officers that provides the invisible glue needed to bind the fragments of the administrative process together in any bureaucracy. Such cohesion is typically sustained by the kind of spontaneous, open collaboration among peers that was inadvertently suppressed by the zero sum characteristics of the controlled rating system.

Resource Utilization

While the forced distribution requirement did produce the desired level of differentiation and a "de facto" deflation of ratings, it also generated some untoward side effects in the areas of job assignments and resource utilization. As evaluation reports containing the controlled ratings accumulated in personnel management records, an unofficial caste system began to emerge, effectively dividing the officer corps into two groups: the "haves" (above average), and the "have nots" (average and below). Prospects for mobility between the two groups were substantially limited by a job assignment process that relied heavily on the aggregate assessment of potential recorded in an individual's performance evaluation file. The "have nots" were faced with an almost insoluble "Catch 22" situation. To improve their evaluations, and chances for promotion, they had to demonstrate increased potential in more demanding jobs of increased responsibility. However, these were the jobs routinely reserved for those who had already received "above average" ratings of potential. Individuals not in that category were usually eliminated from consideration. The net result was the development of a closed system in which 50% of the officer corps understandably felt that, regardless of their efforts, they would remain outside the winners circle.²⁹

Unintended consequences among the above average "haves" were equally as significant. Effectively identified as individuals on the fast track, they were expected to accept the most demanding, responsible positions, typically located in organizational headquarters and other selectively manned units where talent and ability were at a premium. Within these organizations, the forced distribution of ratings was the same as in any other unit and required a differentiation of potential among a collection of the best talent available. The individuals selected to fill these priority positions,

based on their previous above average ratings, realized that if they accepted such assignments, 50% of them would no longer be counted among the "haves" after the next rating cycle. To many, it made more sense to look for a position in a unit of lesser priority, staffed with a greater number of "have nots" where one could compete as a "big fish in a little pond" and stand a better chance of getting above average ratings. As a result of such calculus, many outstanding individuals tended to resist assignments to selectively manned units where their talents could best be utilized, increasing the risk of malassignments and generally making the optimum use of human resources more difficult.³⁰

CLOSING OBSERVATIONS

The preceding paragraphs provide few prescriptive remedies for the difficulties inherent in any evaluation process. They do tend to reinforce many of the observations in the literature concerning the unintended consequences of appraisal systems that are insensitive to the dynamics of human organization. The experiences described may be colored somewhat by the character of the military environment, but they should nonetheless be instructive for those concerned with the assessment of individual performance in large organizations.

Among the lessons suggested, it seems clear that evaluation systems requiring a forced distribution of ratings are likely to produce dysfunctional consequences that far outweigh any administrative utility they may afford. Although their conceptual simplicity is attractive, it masks the unavoidable conflict they generate and obscures the hefty price tag they carry in terms of human capital within the organization. The study also lends support to the now familiar caution against the use of composite ratings that ostensibly capture an individual's relative, total worth to the organization. Such single, global measures have an aura of authenticity about them, but painfully ignore the fact that successful organizations require a "carefully coordinated set of widely different technical and social skills in response to the particular task at hand."³¹ The present analysis would indicate that attempts to devise comparative, aggregate measures of these essential contributions remain spurious at best. Similarly, it attests to the futility of trying to devise a single performance evaluation instrument to serve all of management's needs. The annual appraisal report adopted by the Air Force, was designed to provide assessments of both performance and potential for use in job assignments, training decisions, promotions, and retention. In retrospect, this multiplicity of functions placed unrealistic demands on the entire process and undermined its basic credibility, making it of questionable value in a total management context.

Perhaps more importantly, this study provides an illustration of one of the more common pitfalls of public personnel administration. All too often, personnel technicians succumb to a kind of myopic suboptimization that values administrative neatness and standardization above the broader objectives of organizational sustainment and development. Consequently, their initiatives, though well intended, are frequently viewed by managers more as impositions to be tolerated than as constructive

contributions to the success of the organization. The appraisal system discussed here, reflected this kind of goal inversion and was broadly rejected as an intrusion by the "personnel people", that disrupted the harmony of the organization and made the real task of management that much more difficult. The procedures implemented here, for the most part, were insensitive to the interrelatedness of the social and technical components that constitute any personnel management system and focused almost exclusively on the latter. They ignored the fact that "organizational reality is a dense fabric of technical, social, and psychological threads, woven together in interlocking patterns."³³ To so tamper with a single thread, without due regard to the continuity of the fabric as a whole, invites predictably unfortunate consequences in terms of organizational health and vitality.

NOTES

¹See Alan L. Patz, "Performance Appraisal: Useful But Still Resisted", *Harvard Business Review*, vol 53 (May/June 1975), pp. 74-80. Robert Lazer and Walter Wilkstrom, "Appraising Managerial Performance: Current Practices and Future Directions", *Conference Board Report No. 723*, (New York: The Conference Board Inc. 1977), as well as "Performance Appraisal Programs", *Personnel Policies Forum Survey No. 135*, (Washington, D.C.: Bureau of National Affairs Inc., Feb. 1983).

²Some of these dysfunctions have been described by V.F. Ridgway in "Dysfunctional Consequences of Performance Measurements", *Administrative Science Quarterly*, Vol 1 (Sept 1965), pp. 240-247., D. McGregor, "An Uneasy Look at Performance Appraisal", *Harvard Business Review*, Vol 35, (May/June 1957), pp. 89-94., P. Thompson and G. Dalton, "Performance Appraisal: Managers Beware", *Harvard Business Review*, Vol 48, (Jan/Feb 1970), pp. 149-157., J. Nalbandian, "Performance Appraisal: If Only People Were Not Involved" *Public Administration Review*, Vol 41, (May/June 1981), pp. 392-96.

³W. Sayre, "The Triumph of Technique Over Purpose" *Public Administration Review*, (Spring 1984), pp. 134-137.

⁴A point made by N. Winstanley, "How Accurate Are Performance Appraisals?", *The Personnel Administrator*, Vol 25, (Aug 1980), pp. 55-58.

⁵This article is based in part on research done by the author in 1977 while a member of a Department of the Air Force special task force charged to assess, among other things, the organizational impacts of the system used to evaluate Air Force Officers from 1974 to 1978. The research included interviews with over 200 officers evaluated under the system, participation in several conferences on the subject, and a review of relevant internal Air Force documents, surveys, and reports.

⁶These are described in various editions of *Air Force Manual 36-10* and *Air Force Regulation 36-10* dealing with officer performance evaluation and training reports.

⁷Most personnel administration texts inventory these approaches. See for example O. Glenn Stahl, *Public Personnel Administration*, (New York: Harper & Row, 1983), D. Klingner, *Public Personnel Management: Contexts and Strategies*, (Englewood Cliffs N.J.: Prentice Hall 1980), and G. Siegel & R. Myrtle, *Public Personnel Administration: Concepts and Practices*, (Boston: Houghton Mifflin Co., 1985).

⁸Between 1958 and 1981, there were eight different editions of the AF Form 707, "Officer Effectiveness Report".

⁹J. Shafritz et al, *Personnel Management In Government*, (New York: Marcel Dekker Inc., 1986) and D. Klingner & J. Nalbandian, *Public Personnel Management: Contexts and Strategies*, (Englewood Cliffs N.J.: Prentice Hall, 1985) provide representative descriptions of objectives to be served by performance appraisal systems.

¹⁰See O. Glen Stahl, *op. cit.*, p. 264

¹¹E. Yager, "A Critique of Performance Appraisal Systems", *Personnel Journal*, Vol. 60, (Feb 1981), p. 130., Stanly Sokolick, "Guidelines In Search of Effective Appraisals", *Personnel Journal*, Vol. 46 (Nov

1967), pp. 663., and P. Thompson and G. Dalton, "Performance Appraisal: Managers Beware", *Harvard Business Review*, Vol. 48 (Jan/Feb 1970), pp. 155-57., all raise this caution.

¹²Actually, the Air Force Human Resources Laboratory at Brooks Air Force Base, San Antonio, Texas, had been exploring alternative system designs since 1961. Their efforts, along with those of several other study groups, were accelerated and brought together in the early 1970s.

¹³D. McGregor, *op. cit.*, p. 90, and J. Nalbandian, *op.cit.* p. 394 describe this tendency to avoid negative evaluations.

¹⁴C. Tyer, "Employee Performance Appraisal: A Process in Search of a Technique", in S. Hayes and R. Kearney, *Public Personnel Administration: Problems and Prospects*, (Englewood Cliffs, N.J.: Prentice Hall 1983), p. 120, describes this tendency of public organizations to narrowly focus on this objective.

¹⁵See Siegel and Myrtle, *op. cit.*, p. 318 for a good discussion of forced distributions.

¹⁶*Ibid.*

¹⁷L.L. Cummings and D. Schwab, *Performance In Organizations: Determinants and Appraisal*, (Glenview: Scott Foresman & Co. 1973), pp. 83-85, and K. Duffy and R. Webber, "On Relative Rating Systems", *Personnel Psychology*, Vol. 27, (1974) pp. 307-311 provide some examples.

¹⁸O.G. Stahl, *op. cit.*, p. 264.

¹⁹L. Cummings and D. Schwab, *op. cit.*, offer a warning about the application of forced distributions to such organizations.

²⁰E. Yager, *op. cit.*, p. 132.

²¹See A. Maslow, *Motivation and Personality*, (New York: Harper and Row, 1954), and F. Herzberg et al, *The Motivation to Work*, (New York: John Wiley & Sons, 1969) for discussions of these factors.

²²See for example Rensis Likert, "Motivational Approaches to Management Development", *Harvard Business Review*, (Jul/Aug, 1959), p. 75, H. Meyer et al, "Split Roles in Performance Appraisal", *Harvard Business Review*, (Jan/Feb, 1965) pp. 123-129, and Thompson and Dalton, *op. cit.* pp. 153-155.

²³C. Levine and G. Wolohojian, "Retrenchment and Human Resources Management: Combating the Effects of Uncertainty", in Hayes and Kearney, *op. cit.*, pp. 175-188.

²⁴C. Barnard, *The Functions of the Executive*, (Cambridge: Harvard University Press, 1938) discusses the importance of the balance between organizational inducements and personal commitments.

²⁵J. Conant, "The Performance Appraisal: A Critique and an Alternative", *Business Horizons*, Vol 16, (Jun 1973), p. 77 and R. Likert, *op. cit.*, p. 75 both discuss the impacts of appraisals on supervisor/subordinate relationships.

²⁶Nalbandian, *op. cit.*, p. 394.

²⁷A tendency described by E. Yager, *op. cit.*, p. 130.

²⁸See F. Thayer, "Civil Service Reform and Performance Appraisal: A Policy Disaster," *Public Personnel Management*, Vol. 10, (Symposium, 1981), pp. 20-28.

²⁹Thompson and Dalton, *op. cit.*, p. 155.

³⁰R.E. Jeffers, "The Influence of 'Gaming' on the Current OER System", *Unpublished Research Report*, (Maxwell AFB, Alabama: Air Command and Staff College, Air University, 1974), details non-productive behavior patterns resulting from the forced distribution.

³¹Thompson and Dalton, *op. cit.*, p. 154.

³²D. Brinkerhoff and M. Rosenbeth, "Appraising the Performance of Performance Appraisal", *Sloan Management Review*, Vol 21, (Spring 1980), p. 4.

Copyright of Public Personnel Management is the property of International Public Management Association for Human Resources and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.