



FORCED RANKING AND FORCED DISTRIBUTION

By Dick Grote

An excerpt from the forthcoming book
Forced Ranking: Making Performance Management Work
To be published by the Harvard Business School Press in October 2005

There is little argument that those who contribute the most, have the greatest potential, and shoulder more than their share of the load, should be more generously rewarded. Likewise, almost no one would argue that those who contribute the least should be encouraged to either increase the quality of their contribution, or find employment at another organization where their talents and contributions are more in line with the organization's expectations. "Executives owe it to the organization and to their fellow workers not to tolerate nonperforming individuals in important jobs," notes management authority Peter Drucker. "It may not be the employees' fault that they are underperforming, but even so, they have to be removed."

The challenge — how do you figure out who's in which group? If conventional performance appraisal systems were actually used as they were designed to be used, the need for any more rigorous procedure would be minimal. Most appraisal systems ask managers to identify their distinguished performers, their solid Steady-Eddies, and their also-rans. But, as we've seen, too many managers would rather fudge

the facts and inflate their ratings than to face having the hard conversations that truthful performance assessments often require.

The familiar performance appraisal system will fail if managers set their goals soft and have performance expectations that are low. To drive rigor and truth into performance appraisal assessments and discussions, many organizations today employ a technique called *forced distribution*, the setting of maximums and minimums in performance appraisal ratings to assure that there is differentiation.

One of the most common misunderstandings that arises when people talk about performance management involves the difference between forced ranking and forced distribution. While most people in organizations are familiar with the term, "forced ranking", it turns out that they use the phrase to describe two very different processes. In some cases they are referring to the procedure made famous by Jack Welch and General Electric, in which managers are required to evaluate their employees and assign them into

pre-determined categories based on their performance, potential and promotability. In other cases they use the "forced ranking" term to refer to a requirement in the company's performance appraisal process that allows only a certain percentage of employees to get the highest rating and requires that some minimum number must be assigned to the lowest categories of performance. And sometimes they mean something else entirely.

Here's a personal example of the multiple-meaning confusion I've described: Within the same month recently I was invited to spend a day providing an executive overview to the senior leadership teams of two large corporations on the subject of forced ranking. My assignment, as I understood it (and my clients seemed to be very clear in spelling out their expectations for the day as we discussed them in our telephone conversations in advance of the engagement), was to discuss the use of the forced ranking procedure that my new book, *Forced Ranking: Making Performance Management Work*, will describe. From our conversations it was clear that my job was to describe the ways in which companies like Sun and PepsiCo and GE were using the forced ranking technique of assigning people into buckets (like top 20, vital 70, bottom 10) based on performance and potential, analyze the pros and cons of the approach, and then in smaller sessions with their senior HR managers review their own forced ranking process with the goal of fine-tuning it.

That's what I thought. But in both cases when I arrived on the scene it turned out that there was a significant amount of confusion and misunderstanding — on my part, on the part of the senior executive who engaged me, and on the part of the members of the executive team I was to address — about precisely what we were talking about. In the first organization, an entertainment colossus, the organization's objective was actually to tune up its performance appraisal procedures. In doing so, it was considering giving managers some fairly rigid guidelines on what

the company's top executives felt an appropriate distribution of appraisal ratings would look like. In the second case, a large telecommunications firm, the company for the first time in several years had a few extra dollars it could put into a bonus pool. They were considering using some kind of quick totem-pole stack-ranking process to make the best decisions about who should get bonuses and how the dollars should be allocated.

With both organizations the day I spent with them ended up being highly productive. All of their goals got achieved. Both times, however, there was a short period at the start of the day awkwardly sorting out exactly what we were talking about when we used the same "forced ranking" term to mean quite different things.

Forced ranking and forced distribution are two different processes. "Forced ranking," the system GE made famous but that is used in a multitude of other well-managed, tough-minded corporations, will be the subject of my presentation to the North Texas Compensation Association meeting on June 16, 2005. In this paper, though, I'll discuss the forced distribution of appraisal ratings and examine how this technique can be used effectively in a traditional performance appraisal system.

How Forced Distribution Works

Forced distribution is a procedure that involves tweaking the rules of the performance appraisal system to either request or require a certain distribution of performance appraisal ratings. Consider a company that has a five-level rating scale — Distinguished / Superior / Fully Successful / Needs Improvement / Unsatisfactory. Unhappy with the fact that every year almost everyone in the firm is rated in the top two categories, senior management decides to adopt a policy that restricts the number of people who can get the top rating and requires a certain amount of lower level ratings. Their forced-distribution scheme might look like this:

Distinguished	5%
Superior	20%
Fully Successful	50%
Needs Improvement	20%
Unsatisfactory	5%

If rigidly enforced, the new policy will certainly assure that there is differentiation in performance appraisal ratings. But two major problems immediately arise with the proposed distribution scheme above. First, there is no flexibility in the percentages of ratings allowed. As a result, the manager of a low-performing group is required to assign five percent of her subordinates into the "Distinguished" category and another twenty percent into the "Superior" category when, in truth, hardly any of them deserve even a "Fully Successful" rating. Likewise, the manager of a group of all-stars is restricted from giving appropriately high ratings to his outstanding cadre of performers. And exactly half of the people in each group, regardless of either the relative or absolute quality of their performance, will be rated as Fully Successful.

Another problem with the proposed distribution scheme is that it exactly mirrors a bell-shaped curve. Regardless of how well people in the organization actually perform, in this example the manager is required to allocate performance ratings to fit a Gaussian statistical model. But a pure bell curve is almost never accurate in describing the distribution of human performance. For an legitimate bell-shaped Gaussian curve to be valid, two conditions must be present. First, there must be a sufficiently large population under assessment. While the company as a whole may have enough people to support the math behind a bell-shaped curve, individual work units of four or eight or a dozen employees just aren't large enough for a bell curve to be valid.

More important, in addition to a sufficiently large population, there must be a *random distribution* of the elements that make up that population for a bell curve to be valid. But the employee population of an organization is never the result of random distribution. Companies don't hire people at random, selecting say every 14th applicant. Instead, they sift the applicant pool carefully to identify the most promising. Companies don't promote people at random, selecting people for advancement by lottery or on an alphabetical basis. Instead, they promote the best talent they've got. Almost every company provides training programs and developmental experiences to enhance the quality of talent and performance throughout the organization, thus further de-randomizing the population. And in spite of the complaints that they don't do it fast enough or soon enough, organizations still weed out and terminate some members of the workforce whose performance is simply unacceptable.

Thus a pure bell-shaped curve is almost never an appropriate model for the distribution of scores in a performance appraisal system. The distribution must take into account the fact that there will always be a positive skew — an appropriate positive skew — caused by organizational efforts to improve the quality of the employee population.

A better approach might be to allow some flexibility in the allocation of ratings, and allow a positive ratings skew to reflect the reality that the members of the organization are not performing at random. Upon reflection, our hypothetical organization that wants to assure a degree of differentiation in its performance appraisal results might alter its first draft of the proposed distribution scheme and propose a model that more accurately reflects the reality of human performance:

Distinguished	5% maximum
Superior	20 - 30%
Good Solid Performer	50 - 60%
Needs Improvement	10 - 15%
Unsatisfactory	2 - 5%

Now, instead of requiring that exactly five percent of all employees be rated Distinguished regardless of the actual level of their performance, the revised distribution scheme allows managers to place up to five percent in that category — but there's nothing that says that anyone must be rated Distinguished. At the other end of the scale, the requirement has also been adjusted to provide a range of two to five percent of employees who are expected to fall into the Unsatisfactory category. There must still be some Unsatisfactory appraisal ratings assigned but the requirement has been significantly reduced.

Instead of the rigid and fixed percentages that were required for each of the three middle performance appraisal ratings in the initial model, ranges have now been provided. This will increase the flexibility of the system and allow managers of particularly poor-performing or high-performing work units to reflect that fact in the appraisal ratings they assign to their team members. Finally, the rigid bell-curve requirement has been adjusted to allow twice as many people to be assigned to the rating category one step higher than the middle rating than the one a step lower than the middle. In other words, twenty to thirty percent of all employees can expect to receive a Superior rating while only ten to fifteen percent of all employees — half as many — are expected to fall into the Needs Improvement category.

But some important questions still need to be addressed:

- Is the distribution scheme described above in reality appropriate for an actual organization?
- How well will the system described actually meet the needs of the two managers described earlier, one who is saddled with a bunch of goof-offs and goldbricks, the other with a cadre of champions?
- How should it be applied to a work unit with a very small population? For example, if there are only three people in the department, does that mean that no one can be rated Distinguished?

- Should the distribution percentages merely be suggested, with flexibility allowed for individual managers to vary from the ranges if they can present a compelling case for variance, or must the distribution requirement absolutely be met by all?
- When, if ever, should exceptions be allowed? If a manager can demonstrate through hard business results that her team has outperformed any comparable unit in the organization, should she be allowed to award a greater percentage of higher ratings? And if another manager's unit produces only mediocre results, should he be prevented from awarding a Distinguished or Superior rating to anyone on the team?
- How do you control for variations in the standards and expectations of various managers? In one department the manager may evaluate a subordinate with a 98% customer satisfaction score as Needs Improvement; in the next department over the same 98% customer satisfaction score may earn the employee a Superior rating.
- Who will police the system, making sure that the guidelines are followed? Is that line management's job? Is it HR's? And what happens if the policeman is asleep on the beat and allows variances to get by?
- What will happen when a manager (or a whole department) ignores the guidelines? Will managers have to take their appraisals back and assign a different rating (provoking morale-destroying, buck-passing whines from spineless managers to their subordinates that "I really rated you Superior but personnel forced me to lower it to Fully Successful.") And what will happen if the manager has already had the performance appraisal discussion with the individual when the variance from the guidelines is noticed? Will she still have to go back and change it?

Forced distribution requirements can have an appropriate place in the design of performance appraisal systems, particularly when they are structured in such a way as to recognize that in any work group there are likely to be many more good performers than poor ones, and they provide a mechanism for managers of exceptional work teams to secure an exemption from the guidelines' requirements through demonstrating genuine excellence in business results.

If organizations create a set of distribution guidelines that combine flexibility with a tough-minded demand for differentiation, and make sure that appropriate answers are found for each of the questions I've raised above, then a forced distribution of performance appraisal ratings can help drive the truth into the performance management process.

Dick Grote is Chairman and CEO of Grote Consulting Corporation in Dallas, Texas. He is the developer of GROTEAPPROACH, the web-based performance appraisal system, and the author of the books, *Discipline Without Punishment* and *The Complete Guide to Performance Appraisal*. Both books were major book club selections and have been translated into Chinese and Arabic.

Paramount Pictures bought the movie rights to *Discipline Without Punishment* and produced the video series "Respect and Responsibility" with Dick as host. His latest book, *The Performance Appraisal Question and Answer Book*, was published in May, 2002. His new book, *Forced Ranking: Making Performance Management Work*, will be published by the Harvard Business School Press in October 2005.

For five years, Dick was a regular commentator on National Public Radio's "Morning Edition" program. His articles have appeared in the *Harvard Business Review* and *The Wall Street Journal*. His biography appears in *Who's Who in America*.

Copyright © Dick Grote 2005