Forced Ranking
Time to dismiss this underperformer?

Andrew MacLennan explains...

The assessment of individual employee performance has long been a tricky art. It is dogged by problems including assessor subjectivity, the influence of negotiation skills and low line manager commitment to the task. For many organisations, the emergence of forced distribution ranking systems seemed like a silver bullet. It promised to make managers think about performance criteria and differentiate employee performance, separating the wheat from the chaff. High performers could be exalted and rewarded. Underperformers could be pinpointed and dealt with, via remedial support and if necessary, dismissal.

The practice was of course made famous – and perhaps palatable – by Jack Welch, who as CEO of General Electric went so far as to fire the poorest performing ten percent of his workforce[1]. Three McKinsey consultants, in their book The War for Talent, further fuelled interest and belief in 'vitality curves' and 'rank and yank' systems[2]. Other theorists continue to champion the approach, highlighting the importance of not only focusing on top performers but also grasping the nettle with underperformers.

Many organisations have also been attracted to forced rankings by the apparent financial control offered. Defining performance assessment patterns offers greater control over related rewards (such as salary increases and bonus distributions) and allows some advance signalling of the rewards on offer for specific performance levels.

In the UK, forced ranking systems are growing in popularity. According to a recent survey, 45 percent of managers are required to rank employees' performance using a pre-determined distribution[3]. However, dark clouds are looming for the practice and doubts are being cast over its effectiveness.

In the United States, there are growing legal concerns. As long ago as the mid-seventies, Sandia lost an unfair dismissal case on the basis of age discrimination, being unable to show that their forced ranking system produced a valid assessment of individual performance[4]. More recently class action lawsuits have reportedly been filed against Capital One, Conoco, Ford, General Motors, Goodyear and Microsoft, for various types of alleged discrimination[5]. Some of these firms and many others in the US have since abandoned their 'rank and yank' systems.

Aside from the legal issues, questions are being asked about whether forced ranking actually works as intended. A recent survey of HR professionals suggested there are widespread concerns about the practice[6]. Another survey of US firms conducted by Ed Lawler suggests that the use of forced ranking systems is associated with lower effectiveness of performance appraisals – particularly if routine dismissals follow[7].

A small proportion of firms do go so far as to dismiss a predetermined proportion of their lowest performing employees. Proponents of this approach excitedly point to a recent academic study, which suggests that forced rankings and automatic dismissals may produce a 16 percent improvement in "workforce potential" after two years[8]. Dick Grote says it shows, "the basic hypothesis underlying the forced ranking, rank-and-yank methodology is solid"[9]. Business Week described the study as providing "hard data" to support the use of forced rankings[10]. However the study was actually an entirely theoretical simulation, built upon numerous untested assumptions and no real world observations. The researchers themselves carefully caveat, "Like all simulations, ours is an incomplete representation of reality" and "important effects (e.g., on morale, productivity, or profitability) are largely outside the scope." They clearly state they were not seeking to inform whether forced distribution ranking systems were good or bad for organisations. Decision makers should be aware that despite some of the claims, there is currently no sound empirical evidence supporting the use of forced ranking systems.

One issue the theoretical simulation raises is the ongoing benefits from using 'rank and yank' systems. Common sense would suggest that the better organisations are at removing underperformers, the fewer underperformers they will have. A law of diminishing returns must set in and indeed, the simulation suggested that the 16 percent improvement would fall to only two percent after six years. This implies that if 'rank and yank' should be used at all, it should only be used for short

periods and under specific circumstances. It sounds impressive to say that organisations should keep 'raising the bar' but this can ignore the realities of the labour market and hard financial implications of staff turnover. Organisations that keep cutting the bottom five or ten percent of their workforce are in time likely to start dismissing capable employees, with no certainty that their replacements will be any better – let alone so much better as to justify the costs, disruption and productivity losses caused by enforced turnover.

There are other regularly cited problems with the practice. 'Survivor guilt' can affect those who escape workforce culls, and any perceived unfairness in the systems used is likely to invoke anxiety amongst survivors as well as those dismissed. Perceived unfairness in forced ranking systems is very common, and of course affects even those organisations that do not routinely dismiss lowest performers but use ratings to drive development, promotion and reward decisions.

It is a myth that forced ranking systems make assessors more objective. Being forced to differentiate employee performance is not a substitute for improved accuracy of performance assessments. Using rankings is neither a prerequisite to identifying clear assessment criteria (as some suggest[11]), nor does it overcome assessors' tendencies to be influenced by factors that do not relate to job performance. Being fallible humans, managers are prone to give high ratings to people like themselves, people they like, people they hired, good negotiators, those who have performed well very recently and so on, regardless of the mechanism used to record their assessments[12].

Another fundamental problem is that distribution methods never accurately reflect reality. Normal distribution ('bell') curves statistically fit large samples (e.g. many thousands of employees) affected by non-systemic patterns. There is no reason to assume that the performance of 30 or even 300 employees in a department or function will coincide with such patterns. 'Totem pole' systems whereby employees are ranked one-by-one are equally misleading, implying that all employees perform at (measurably) different but evenly-spaced levels.

Serious problems result from distribution patterns being applied to different structural subunits[13]. If a company's marketing function is performing brilliantly and attracts fabulous people but its IT function is a dismal failure, is it appropriate for the good but lowest performers in marketing to be punished and the poor but top performers in IT to be rewarded?

Comparative performance evaluations have another disturbing feature, in that often neither organisations nor employees can tell if individuals' performance is improving, stagnant or declining, from one year to the next. Without absolute performance benchmarks, there is no way to tell – and that throws into doubt efforts made to improve individual performance. This is a particular problem in senior roles and complex ones, where hard performance data (such as sales made or applications processed) is rarely available. This problem is compounded because in these roles the performance gap between the lowest and highest performers is greatest[14]. Using forced distribution rankings simply magnifies this key challenge.

No one has ever suggested organisations regularly tell the vast majority of their people they are average or worse than average. Yet that is exactly what forced ranking systems do.

Ironically, motivation problems are common with forced ranking[15]. Transparency around what rewards will flow from specific performance ratings ticks some of the boxes according to Victor Vroom's famous 'expectancy theory'. However, no one has ever suggested organisations regularly tell the vast majority of their people they are average or worse than average. Yet that is exactly what forced ranking systems do. There is a wealth of evidence that such 'labelling' and beliefs about personal potential are self-fulfilling prophecies[16]. Tell people they are doing well and have bags of potential and they will generally prove you right; tell them they are failing and you'll probably come to be right about that too. Of course, managers don't need sophisticated research to tell them that, which is partly why they prefer to avoid giving poor evaluations.

Forced ranking systems have caused managers to develop numerous tactics to compensate for these deficiencies. Many distance themselves from the process, explaining that they do not like or agree with it, reinforcing resentment over its use. Some managers – particularly those who have to perform regular evaluations – rotate employee evaluations to restore a sense of equity. Others retain serious underperformers for prolonged periods to ensure they have candidates for low ranking come appraisal time. Equally, sometimes underperformers are not dealt with because they fall too high up the ranking and escape attention.

Managers also tend to give similar ratings to most employees because they recognise another danger – that forced rankings can create unhealthy competition between individuals. This belief is supported by empirical research findings[17]. For organisations that require good teamwork and cooperation between people, forced ranking systems that will affect rewards, promotions and even organisational survival are risky choices.

Forced ranking systems also exacerbate the 'judge and helper' dilemma in performance appraisals. They direct the attention of employees to their performance evaluation rather than their development needs. The more significant the implications of a specific ranking, the less likely it becomes that employees will approach managers to discuss skill gaps or request developmental support – doing so may simply make them more likely candidates for the 'underperformer' label.

There are signs that practitioners are recognising the extent of these problems.

Last year Microsoft abandoned its forced ranking system, wanting to encourage more open conversations between managers and employees about performance and boost morale. At the same time, it introduced a wide range of new benefits oriented to make its employees' busy lives a little easier. Even General Electric, in the wake of Jack Welch's departure, has softened its use of forced rankings. Managers no longer have to stick rigidly to the 20-70-10 percentage split introduced by Welch (the bottom ten percent being termination candidates), and underperformers are given considerably more support than previously[18].

The greatest problems with forced ranking systems perhaps lie in the fundamental assumptions that underpin their use. They focus squarely on trying to improve organisational performance through increasing the performance of the individuals in organisations, be it through motivation, development or enforced turnover. As we have seen, many of these mechanisms may be flawed. But is the focus on individual capability the route to organisational success?

There is evidence that high performing individuals are more likely to create high performing teams and that individual talent does of course affect performance in a role – but only up to a point[19]. There is also evidence that the way in which teams are brought together and managed has a significant effect on how they perform. For example, team membership stability is strongly associated with team performance[20]. In fact, we know of numerous other factors affecting the performance of teams and organisations – most of which have very little to do with individual traits. There is thus a great danger with the obsession with individual 'talent'. Talent is not some kind of fixed asset possessed by individuals. Individual performance can change radically over time and talent, if it really exists as a meaningful, measurable quality, is tied up much more closely with organisational context than forced ranking systems imply.

It is fascinating that despite all their limitations, well chosen and applied selection techniques are quite effective at predicting the performance of individuals in organisations[21]. But these tools tell us much more about how individuals will fit into and be judged to have performed once working in the selecting organisations, than they do about individuals' potential to help organisations achieve their objectives. That depends largely on how organisations integrate and align individual contributions.

Behavioural scientists have known for years that we consistently underestimate contextual influences on individual performance and correspondingly give individuals too much credit or blame for their successes and failures[22]. Forced ranking systems do not acknowledge organisational performance constraints and arguably even discourage diagnosis of them. Why would anyone tackle the challenge of problem diagnosis believing that fighting a 'war for talent' is the key to success?

What is the most important lesson from the short but chequered history of forced ranking? Probably that in organisations the clue to success is in the name. The way talent is organised is far more important than individual talent. Our focus should be on how people work together, rather than their systematic segregation.

Andrew MacLennan is Managing Director of Strategy Execution Ltd, which he founded to provide consulting and management development services. He is also a global faculty member of Duke Corporate Education, lectures at Edinburgh Business School and leads the ifs School of Finance strategy execution masterclass series. Andrew is co-author of Making Strategies Work published by Pearson and is currently writing a book on strategy implementation for Routledge. www.strategy-execution.co.uk

1 Welch, J. (2001) Jack: Straight from the gut. New York, NY: Warner Books.
2 Michaels, E., Handfield-Jones, H. & Axelrod, B. (2001) The war for talent. Boston, MA: Harvard Business School Press.
3 Houldsworth, E. & Jirasinghe, D. (2006) Managing and measuring employee performance. London: Kogan Page.
4 Lawler, L.E. (2002) The folly of forced ranking. Strategy & Business, 28, 28-32.
5 Osborne, T. & McCann, L.A. (2004) Forced ranking and age-related employment discrimination. Human Rights Magazine, Spring Issue.
6 Novations Group, Inc. (2004) Uncovering the growing disenchantment with forced ranking performance management systems. Novations Group, Inc.
7 Lawler, L.E. (2003) Reward practices and performance management system effectiveness. Organization Dynamics, 32, 396-404.
8 Scullen, S.E., Bergey, P.K. & Aiman-Smith, L. (2005) Forced distribution rating systems and the improvement of workforce potential: A baseline simulation. Personnel Psychology, 58, 1-32.
9 Grote, R. (2005) Forced ranking: Making performance management work. HBS Working Knowledge: http://hbswk.hbs.edu/archive/5091.html.
10 Anonymous. (2006) The struggle to measure performance. Business Week, 9th January.
11 Grote, R. Op. cit.
12 Asch, S.E. (1946) Forming impressions of personality. Journal of Abnormal and Social Psychology, 41, 258-290; Avery, R.D. & Campion, J.E. (1982) The employment interview: a summary and review of research. Personnel Psychology, 35, 281-322; Cardy, R.L. and Dobbins, G.H. (1986) Affect and appraisal accuracy: Liking as an integral dimension in evaluating performance. Journal of Applied Psychology, 71, 672-678; Lefkowitz, J. (2000) The role of interpersonal affective regard in supervisory performance ratings: A literature review and proposed causal model. Journal of Occupational and Organizational Psychology, 73, 67-85; Miller, N. & Campbell, D.T. (1959) Recency and primacy in persuasion as a function of the timing of speeches and measurements. Journal of Abnormal and Social Psychology, 59, 1-9; Schoorman, F.D. (1998) Escalation bias in performance appraisals: an unintended consequence of supervisor participation in hiring decisions. Journal of Applied Psychology, 73, 58-62.
13 Cummings, L. & Schwab, D. (1973) Performance in organizations: Determinants and appraisal. Glenview, IL: Scott Foresman & Co.
14 Schmidt, F.L. & Hunter, J.E. (1998) The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. Psychological Bulletin, 124, 262-274.
15 McBriatry, M.A. (1984) Performance appraisal: Some unintended consequences. Public Personnel Management, 17, 421-434.
16 e.g. Livingston, J.S. (1969) Pygmalion in management. Havard Business Review, 47, 81-89.
17 Garcia, S.M. & Tor, A. (2007) Rankings, standards, and competition: Task vs. scale comparisons. Organizational Behavior and Human Decision Processes, 102, 95–108; McBriatry, M.A. Op. cit.
18 Anonymous. Op. cit.
19 Locke, E.A., Tirnauer, D., Roberson, Q., Goldman, B., Latham, M.E. & Weldon, E. (2001) The importance of the individual in the age of groupism. In Tuner, M.E. (Ed.), Groups at work. Mahwaw, NJ: Earlbaum, 501-528.
20 Hackman, R. (2002) Leading teams. Boston, MA: Harvard Business School Press; Eisenhardt, K. M. & Schoonhoven, C.B. (1990) Organizational growth: Linking founding team, strategy, environment, and growth among U.S. semiconductor ventures, 1978-1988. Administrative Science Quarterly, 35, 504-529.
21 Schmidt, F.L. & Hunter, J.E. Op. cit.
22 Ross, L. (1977) The intuitive psychologist and his shortcomings: Distortions in the attribution process. In Berkowitz, L. (Ed.), Advances in experimental social psychology, 10, 173–220. New York, NY: Academic Press.