Best Practices for Performance Management

Manju Abraham, Netapp
Rajen Bose, Yahoo
Balu Chaturvedula, Yahoo
Jay Crim, Google
Kuk-Hyun Han, Samsung
Manisha Jain, Google
Ikhlaq Sidhu, UC Berkeley

College of Engineering
University of California, Berkeley

Fung Technical Report No. 2013.10.10
http://www.funginstitute.berkeley.edu/sites/default/files/PerformanceMgmt.pdf

October 10, 2013
The Coleman Fung Institute for Engineering Leadership, launched in January 2010, prepares engineers and scientists – from students to seasoned professionals – with the multidisciplinary skills to lead enterprises of all scales, in industry, government and the nonprofit sector.

Headquartered in UC Berkeley’s College of Engineering and built on the foundation laid by the College’s Center for Entrepreneurship & Technology, the Fung Institute combines leadership coursework in technology innovation and management with intensive study in an area of industry specialization. This integrated knowledge cultivates leaders who can make insightful decisions with the confidence that comes from a synthesized understanding of technological, marketplace and operational implications.

Lee Fleming, Faculty Director, Fung Institute

Advisory Board

Coleman Fung
Founder and Chairman, OpenLink Financial

Charles Giancarlo
Managing Director, Silver Lake Partners

Donald R. Proctor
Senior Vice President, Office of the Chairman and CEO, Cisco

In Sik Rhee
General Partner, Rembrandt Venture Partners

Fung Management

Lee Fleming
Faculty Director

Ikhalq Sidhu
Chief Scientist and CET Faculty Director

Robert Gleeson
Executive Director

Ken Singer
Managing Director, CET

Copyright © 2013, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.
Abstract: Performance management describes a method of enabling people in an organization to perform at their best in order to accomplish goals and objectives that contribute to the success of that organization. Today’s organizations spend large amounts of time and money implementing systems that attempt to direct, assess and develop their employees. While a company often credits its performance management system as contributing greatly to its success, the same set of processes at another company can result in strong internal dissent or even legal action. This paper highlights the wide range of successes and failures of several performance management practices, and argues for the need for thoughtful design and continual evaluation of any system to ensure that it is in fact helping and not hurting an organization.
Introduction: Performance Management

Performance management is a set of processes focused on maximizing an organization’s performance by ensuring that each employee performs his or her assigned role as well as possible. It attempts to both align an organization's efforts with its goals, and to assess to what extent each employee contributes to achieving these goals. For the employees that need it, these systems often also offer development or remediation opportunities.

Along with effective hiring and compensation policies, good performance management is vital to the success of today’s organizations. The importance of successful performance management is highlighted in a 2002 Corporate Executive Board study, which estimates that various techniques can improve or degrade an organization’s performance by 40 percent.

Many reports suggest that today’s performance management systems are in need of dramatic improvement. For example, a 2004 Watson Wyatt study found that only 61 percent of employees felt that their performance assessment was accurate, and many fewer felt the system at their company established clear goals (35 percent) or helped them improve (30 percent). Likewise, World at Work reported in 2010 that only 30 percent of employees trusted their company’s performance management system.

In this paper, we describe common elements and strategies used when implementing successful performance management systems and explain why the same practices applied in different companies can have drastically different outcomes. We argue for the importance of continual evaluation of any program and suggest some techniques for understanding if and when a system is not achieving its goals.

By providing an understanding of the goals and building blocks of existing performance management systems and tools for evaluating when a company needs to change, we hope companies can improve on the current state of performance management and better use their employees to achieve organizational goals.
Part A: Common Building Blocks of a Performance Management System

1. Plan

The first crucial component, the process of developing and writing objectives for a specific period of time, must be aligned and fully integrated at the company, the team and the individual levels. This joint effort between manager and employee results in a mutually agreed upon set of goals, expectations and key performance areas for the defined period. This plan also includes development areas to enable addressing areas of improvement and career development.

2. Support and Develop

The process of managers and companies providing support is ongoing. Support includes feedback, guidance, resource support, and development opportunities such as on-the-job experiences, mentoring, training, as well as other activities that allow employees to grow and develop professionally at the company. The employee receives open and transparent feedback to enable him or her to meet or exceed goals.

3. Assess
This is the function of evaluating the employee’s performance in the current job through the review of goals and objectives achieved over the performance period. Regular, ongoing feedback is the greatest predictor of employee success and the most important component of the performance evaluation process. When you continually observe your employees, discuss how they are performing against their goals, and provide them with real-time feedback, it should never come as a surprise to the employees. The assessment of employees happens through a calibration process, a method for comparing employees.

4. Recognize and Reward

Providing appropriate recognition and rewards to employees for their performance is imperative. The most important step in annual review and reward processes is the performance review discussion with directly reporting employees. This discussion should strengthen your partnership with your employee by focusing on performance strengths and presenting development areas as opportunities or strong motivators. Finally, managers relay information regarding employee salary promotions (if any), increases, bonuses and equity awards.

Part B: One Size Does Not Fit All

While decades of research have contributed to theories of performance management and basic building blocks seem similar across most companies, agreement on how to implement these theories remains elusive. One size clearly does not fit all. We examine two methods for assessing an employee’s performance: one targeted at decision-making, and another focused on growth. We show how the success of the same strategy can vary at different companies. First, we describe the heated argument over whether forced ranking systems improve or degrade performance. We then look into the pros and cons of 360-degree feedback.

1. Forced Ranking

The system of forced ranking, or assigning pre-determined percentages of individuals in an organization to each tier of performance rating, is perhaps one of the most controversial practices in today’s performance management systems. This is especially true when companies use the bottom performance tier to determine which employees are fired or laid off. Some companies attribute high levels of organizational productivity to the practice, while others cite it as a major contributor to the downfall of industry leaders. Numerous lawsuits have attacked it as a cover for some form of illegal discrimination.

Forced ranking was popularized by General Electric, which instituted the practice in the 1980s under then-CEO Jack Welch. Welch oversaw the near doubling of GE’s revenues to $130 billion in only five years in the mid-1990s and attributed this stellar performance to the company’s ability to retain top talent and eliminate poor performers. Every year under Welch, the company bucketed its employees into three categories: the top 20 percent, the next 70 percent, and the bottom 10 percent. GE then
heavily rewarded the top-rated 20 percent of its workforce and eliminated the bottom-rated 10 percent.

Proponents of this system say it helps create a meritocracy and addresses the perpetual problem of inflated performance ratings that can make it hard to distinguish good employees from bad. Forced ranking has also been used as a way to justify compensation and termination decisions, most notably at American International Group. Before AIG’s contentious government bailout in 2008, most employees were rated “above average,” fueling criticisms that the firm rewarded even those who helped bring about its downfall. By instituting a forced ranking system, they hoped to identify, reward and retain top performers without enraging a public that believed they were spending too much on compensation. Under their new system, only 10 percent of employees will be eligible for the largest bonuses.

The practice hasn’t always been as successful as these cases make it appear, and its detractors claim that it often lowers organizational performance rather than improving it. Critics point out that teamwork is essential for groups to work well, but forced ranking can result in reduced collaboration and even backstabbing as employees compete with each other for the top ratings. They also believe the system is fundamentally flawed because it ignores group dynamics. How does one rank team members who each perform distinct but important roles? Furthermore, while some researchers believe that employee performance tends to follow a Pareto distribution, what guarantees that all groups perform at exactly the same level as each other (and at all times)?

Table B-1: Pros and Cons of Forced Ranking

<table>
<thead>
<tr>
<th>Forced Ranking</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pros</strong></td>
<td><strong>Cons</strong></td>
</tr>
<tr>
<td>“People need to know where they stand. Failing to differentiate among employees—and holding on to bottom-tier performers—is actually the cruellest form of management there is.” Former GE CEO Jack Welch</td>
<td>“Staffers were rewarded not just for doing well but for making sure that their colleagues failed.” How Microsoft Lost It’s Mojo: Steve Ballmer and Corporate America’s Most Spectacular Decline, Vanity Fair</td>
</tr>
<tr>
<td>• Helps create a meritocracy that distinguishes good employees from bad employees.</td>
<td>• Leads to increased competition or reduced collaboration.</td>
</tr>
<tr>
<td>• Eliminates across-the-board inflated ratings.</td>
<td>• Doesn’t correctly address varied performance of different groups.</td>
</tr>
<tr>
<td>• Provides justification for compensating top performers.</td>
<td>• Enforced quotas are arbitrary and may eliminate high performers.</td>
</tr>
<tr>
<td>• Incentivizes unethical or illegal behaviors.</td>
<td>• Is costly when paired with forced terminations.</td>
</tr>
<tr>
<td>• Can lead to legal troubles.</td>
<td></td>
</tr>
</tbody>
</table>

Forced ranking is especially controversial when tied with HR initiatives that terminate workers at the bottom tier who do not improve. For example, Ford Motor Co. required the lowest performers to improve within a year or be fired; Synergy gave low performers only a single quarter to change behavior. While this may be a useful technique for removing under-performers in a bloated organization, critics question whether it is a sustainable strategy. Workers that are removed can likely become competent employees in several years. Firing them can be costly as well as reduce the organization’s ability to perform in the short term while it hires and trains replacements. These terminations are counterproductive if the company is already understaffed or if it is unable to find better performers.
Multiple case studies tell us that forced ranking has had a negative influence:

– In 2001, the energy company Enron collapsed amid a widely publicized accounting scandal. Researchers note that forced ranking may have played a part in the downfall, as traders who were under intense pressure to perform inflated results.

– Microsoft Corp instituted the practice of forced ranking in 2006, despite heavy internal opposition. A 2012 Vanity Fair article argues that this was destructive and drove many good employees away, adding to the continued decline of the tech giant. Microsoft was also the target of litigation alleging that its ranking system unfairly discriminated against African American employees.

– Ford ceased the practice after settling a lawsuit for $10.6 million. Former employees have also brought lawsuits against Capital One, 3M, Goodyear and others.

– GE may have suffered some unintended negative consequences of its forced ranking system. Consultant Robert Tomasko contends that the performance-driven culture likely motivated unethical or illegal behavior throughout its history.

Despite widespread criticism of the practice of forced ranking and numerous lawsuits surrounding it, companies continue to employ it as a standard part of performance management. A 2001 Harvard Business School article suggests that 20 percent of firms use forced ranking, while a 2012 Wall Street Journal article, published almost 10 years later, estimates that 60 percent of Fortune 500 companies do so. This may be attributed to the slow economy and the belief that performance management systems today focus as much on termination or other negative consequences as they once did on rewarding top performers.

See Appendix A for more discussion about how companies that still use forced ranking attempt to overcome its downsides.

2. 360-Degree Feedback

360-degree feedback is a process in which a leader receives anonymous feedback from subordinates, peers, supervisors and customers. First, the person to be rated describes and scores his or her own performance in a self-assessment, after which they receive similar feedback from various other sources. Once all feedback is collected, the rated person discusses it with a coach and attempts to identify both strengths and areas for improvement, putting together an action plan to address the latter.

This method of feedback has gained rapid acceptance in the corporate world. It has been implemented at AT&T, Mass Mutual Insurance, Bank of America, Digital, Tenneco, Amoco, Exxon, General Electric, TRW, Caterpillar, GTE, Syntex, and Chrysler, among others. By 1994, 22 out of Fortune’s 32 most admired companies were using upward or 360-degree feedback and by 1996, 360-degree programs had became nearly universal among Fortune 500 companies, which spend hundreds of millions of dollars annually to support them.

Proponents of 360-degree feedback believe it overcomes many of the shortcomings of traditional, top-down performance reviews from an employee’s immediate supervisor. For example, because it provides multiple perspectives, it can not only overcome biases of a single rater but also better describes how the rated person performs in different roles (as a manager, a peer, etc.) that a single rater may not be able to see. Allowing the rated person to evaluate him or herself provides the unique
opportunity for individuals to understand how their views of their own performance diverge from others opinions.

Not only does 360-degree feedback help employees identify areas for development and improve them, it can have longer-term effects of encouraging additional communications and feedback-seeking activities in the future. Because the rated person is intimately involved in the process by writing a self-assessment, nominating reviewers, and crafting an action plan, this method can improve their trust in the results. Given all of these factors, the Corporate Leadership Council suggests that 360-degree feedback can improve performance by up to nine percent.

Anonymity of raters is a key feature of many 360-degree feedback systems. These systems are built on the premise that people are bad at having open, frank conversations about performance with each other - especially when evaluating a direct supervisor. Many employees may fear retribution for remarks that aren’t positive. A 1997 study investigating the effects of these feelings on a rater’s behavior showed that scores went up significantly when the rater wasn’t anonymous. The claim, then, is that processes like 360-degree that use anonymous raters result in more honest feedback.

While 360-degree feedback has become quite popular for the reasons cited, critics argue that it does not actually achieve its goal of improving employee performance and that its basic tenets are fundamentally flawed. They point out that though multiple raters are likely to give a more thorough assessment of different areas of performance, they also tend to disagree with each other. One study found that raters assessing performance only agree about half of the time, which can confuse the rated person and make it hard to develop an action plan.

### Table B-2: Pros and Cons of 360-Degree Feedback

<table>
<thead>
<tr>
<th>360-Degree Feedback</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
</table>
| **Pros**            | • Offers multiple perspectives on an employee’s performance.  
                     • Self-assessment allows for comparison between raters’ and the rated persons’ perceptions. | • Conflicting opinions of raters can be confusing. |
|                     | • Less susceptible to the biases of a single, top-down rater. | • Rated people are not held accountable for their ratings, which are still subject to biases. |
|                     | • Since the feedback is anonymous, ratings are less likely to be inflated. | • “More negative” ratings are not necessarily “more valid”.  
                     • Anonymous negative feedback can hurt morale without providing enough specifics for improvement. |
|                     | • Identifies areas for development and ways to improve them. | • Many managers do not follow up on action plans that result from 360-degree feedback. |

Anonymity protects raters from retribution for giving honest feedback, but this comes with its own problems. While studies have shown that anonymous feedback tends to be more negative, there’s no guarantee that equates to greater honesty. Raters still have their own agendas and are free to pursue them since anonymity means they cannot be held accountable for their comments or ratings. This can often result in feedback that is both negative and unexpected by the person being rated. One study estimates that 15 to 20 percent of 360-degree feedback falls into this category. Negative, unexpected feedback hurts morale and can damage the credibility of the process unless it’s accompanied by very specific, actionable feedback. When it’s not, the anonymous nature of the
process prevents rated persons from discovering this information and leaves them wondering how to improve.

Many organizations also fail to effectively follow up on action plans created as part of the process. One survey found that 72 percent of those who received 360-degree assessments said their managers did not follow up on action plans, and a larger number (87 percent) felt their efforts to improve were not recognized. This may be because rated persons are confused by the feedback they receive, or because they have no motivation to change. Whatever the reason, raters are less likely to participate in future rounds if they do not see changes as a result of the feedback process.

To address rated persons’ lack of motivation to change, some companies have tried making rated persons accountable for improving after receiving 360-degree feedback. In particular, IBM tried linking lack of improvement to negative action: if managers didn’t remedy poor behaviors within a period of a few years, they would be asked to step back into a purely technical role. This can backfire, however: Research has demonstrated that when ratings become evaluative rather than purely developmental, as many as 35 percent of raters change their ratings. IBM learned this the hard way. They stopped 360-degree feedback programs because the ratings became politically charged, inflated and unreliable.

In fact, a study found that 50% companies surveyed in 1997 that had implemented 360-degree feedback for appraisal had removed it because of negative attitudes from employees and inflated ratings.

For details on how the remaining 50% of companies attempt to make their systems successful, see Appendix B.

Part C: Evaluating a Performance Management System

Audio quid ueteres olim moneatis amici, "pone seram, cohibe." Sed quis custodiet ipsos custodes? Cauta est et ab illis incipit uxor.
- From the Satires of Juvenal, bold text roughly translated: “Who watches the watchers?”

In the sections above, we’ve offered some examples of practices that work well at one company but have disastrous results at another. Despite the best of intentions and a lot of upfront research on best practices, a company may very well implement a system that doesn’t achieve its stated goals the first time around. And even those who love their performance management systems may need to change it. For example, GE no longer ranks employees the way it did with Welch at the helm, and many other proponents of forced ranking have abandoned it even as others continue to adopt the practice.

The lesson? It’s important to understand how to evaluate the system used to evaluate your employees and understand when it’s time for a change.

Unfortunately, measuring the efficacy of a performance management system can in itself be a difficult task. The goal of performance management is to maximize organizational performance, but many other factors are likely to influence the success of a company. Just because an organization has a great quarter or year does not mean that its management practices deserve all of the credit.

Take GE for example. Then-CEO Welch attributed a large part of the company’s success in the 1980s and 1990s to its performance management system. However, one article points out that these techniques were adopted after GE was already an industry leader, so it is difficult to claim
empirically that the techniques were responsible for the company’s continued success. In addition, some attribute GE’s profits during those years to aggressive lobbying and creative accounting that allowed it to avoid almost all taxes in the United States. Furthermore, their returns were likely influenced at least to some extent by the strong U.S. economy during those decades. The entire Dow average was up 1000 percent between the late 1970s and late 1990s, not just GE’s stock. So how can we say with any certainty whether the performance management system was working well?

Retention and satisfaction metrics can also reflect how well a company’s performance management system is working. A sub-optimal system can hurt satisfaction, morale, and retention in an organization. But a company’s ability to retain their talent is also linked to many other factors, like compensation and the competitive landscape, so this measure is far from perfect as well.

Given that it is difficult to directly relate organizational performance and retention to performance management practices, organizations often use a survey-based approach to determine whether or not and to what extent their system is successful. Employees grade the performance management system just as the system is used to grade their performance. The questions on these surveys reflect the different objectives the system is designed to achieve. See the table below for examples.

Table C-1: Sample Questions for Evaluating a Performance Management System

<table>
<thead>
<tr>
<th>Objective</th>
<th>Sample Questions*</th>
</tr>
</thead>
</table>
| Goal-setting   | • Does the program help establish clear goals?  
                 | • Is the performance review system clearly linked to the organizational strategy? |
| Development    | • Has the program helped improve your performance?  
                 | • Did the most recent review produce a development plan for you?  
                 | • Does the program help poor performers improve?  
                 | • Does the program do a good job of providing honest feedback?  
                 | • Do you get enough support from your peers in your initiatives? |
| Assessment     | • Did your most recent evaluation accurately reflect your performance?  
                 | • Does your organization accurately rank people in order of their performance from best to worst? |
| Satisfaction / Retention | • Do you look forward to coming to work every day?  
                            | • Does your work make a difference?  
                            | • Are you happy with the company’s culture? |

Sources:  
*Some questions from the 2004 Watson Wyatt study and the 2002 Corporate Executive Board study

The United States Office of Personnel Management (OPM) suggests that companies also track compliance with the requirements of its performance management system. What percentage of employees participated? Were goals set and reviews submitted on time? If surveys about a system indicate that it is not doing what it was intended to do, it is important to know whether it failed as a result of its design or simply because employees are not doing what’s required of them. The two situations require very different solutions.

OPM also mentions the importance of fairness when evaluating a performance management system. Not only should an organization use surveys to determine whether its employees believe the system is fair, it should also collect and analyze statistical data about performance management outcomes and ensure that they are evenly distributed across gender, race, national origin, age, etc. No matter what employees believe about a system, the system can hardly be considered fair if it systematically treats members of one of these groups differently than those of another.
Part D: Conclusions

Companies have approached performance management with various methods of ranking and feedback gathering and have gotten in return varying degrees of success. Different approaches seem to be relevant based on the culture, values, size and maturity of the company. But fundamental to the success of performance management systems is trust in the organization and open culture. It is also important to not mimic a solution, but design it intentionally, pursue it gradually and measure its effectiveness frequently.
Appendix A: Avoiding the Pitfalls of Forced Ranking

If the practice of forced ranking is here to stay, a literature review offers some suggestions for implementing and managing such a system in as productive a manner as possible. First, companies should only pursue forced ranking if distinguishing the top performers from the bottom is in line with their mission and strategy. For example, it would make little sense for Walmart stores to institute a forced ranking program as the cost of implementing and using such a program would be contrary to their core value of keeping costs low. Likewise, very small companies or others that cannot afford to lose employees or to pay their top talent extra would see little benefit from such a system.

For large companies like GE that strive to offer innovative products or services, being able to accurately identify the top performers in order to develop and retain them is more likely to prove beneficial. Many of the pitfalls of forced ranking can be avoided by focusing only on the top performers instead of using the system to initiate quota-based terminations. In fact, while GE continues to rank employees, it no longer forces out the bottom 10 percent each year. Instead, GE spends its effort on the best employees. A.I.G. operates similarly. However, there is still a danger in fixed quotas: LendingTree initially used its quota system to identify its best, but when the mortgage market soured, those at the lowest tier were the first laid off. Due to the forced distribution, the lowest tier potentially included competent workers. GlaxoSmithKline plc ensures that this will not happen by only using quotas at the top. The top 25 percent of employees compose the “top,” but there is no forced distribution for other employees. An employee only receives the bottom score if their manager believes they deserve it.

When companies do decide to use a forced distribution rating as a precursor to termination, experts - including HR consultant Dick Grote, who has helped more than 50 firms institute such a system - suggest that it be used only as a short-term fix to cut “slacklers” from an organization. By firing the bottom tier only for the first couple years, companies may be able to eliminate employees who do not carry their weight while avoiding the eventual removal of competent workers who fall victim to the quota and the high costs of constantly replacing a large percentage of the workforce.

Appendix B: Avoiding the Pitfalls of 360-Degree Feedback

The process of collecting and acting on 360-degree feedback can be quite expensive. When done improperly, it can hurt morale and erode trust in the performance management system. It is important to measure the effectiveness and quality of the feedback data regularly. Here are some of the reasons companies found that the quality of the feedback was not as high.

Evaluation is a complex task: it requires the rater to gather and analyze someone’s responsibility and impact over a period of time. There is a probability that the raters will simplify the process by forming an overall impression of the rated person rather than focusing on specific behaviors. Raters also have a tendency to focus on more recent interactions since they are easier to recall. Frequent feedback can be used to avoid recency bias.

Raters’ unconscious or conscious psychological bias could also influence the feedback. At the unconscious level, individuals may rate themselves higher than their peers to preserve favorable self-esteem. This in turn may lead to a self-serving bias – the tendency to take personal credit for successful performance and assign responsibility for failure to external causes. It is important that feedback refer to concrete behaviors and recommendations.
Another reason for lower quality data is lack of clarity around role and evaluation criteria, which results in non-normalized, inconsistent feedback. For example, as organizations grow, managers and employees recruited from different companies, mergers and acquisitions, and new offices in new countries, all bring in different expectations. Any time a new cohort is introduced, consistency must be achieved again with adequate training or opportunities to observe and obtain clear, actionable data.

The format of the feedback request also plays a role in quality of the feedback. Quantitative and structured feedback based on common behaviors is easy to acquire, score, and disseminate. However, such data may not have much relevance to a particular workplace and may even yield misleading results. Most of the time, raters do not remember why they scored the way they did, and employees do not remember about their self-assessment.
References


Hill, Andrew. “Forced Ranking is a Relic of an HR Tool.” *Financial Times.* (July 16, 2012)


Terri Linman, “360-degree Feedback: Weighing the Pros and Cons” <http://edweb.sdsu.edu/people/arossett/pie/Interventions/360_1.htm>
The Coleman Fung Institute for Engineering Leadership, launched in January 2010, prepares engineers and scientists – from students to seasoned professionals – with the multidisciplinary skills to lead enterprises of all scales, in industry, government and the nonprofit sector.

Headquartered in UC Berkeley’s College of Engineering and built on the foundation laid by the College’s Center for Entrepreneurship & Technology, the Fung Institute combines leadership coursework in technology innovation and management with intensive study in an area of industry specialization. This integrated knowledge cultivates leaders who can make insightful decisions with the confidence that comes from a synthesized understanding of technological, marketplace and operational implications.

National rankings consistently place UC Berkeley’s undergraduate and graduate programs among the world’s best. Berkeley is home to top scholars in every discipline, accomplished writers and musicians, star athletes, and stellar scientists—all drawn to this public university by its rich opportunities for groundbreaking research, innovative thinking and creativity, and service to society.