Ahead of the curve: The future of performance management

What happens after companies jettison traditional year-end evaluations?

by Boris Ewenstein, Bryan Hancock, and Asmus Komm

The worst-kept secret in companies has long been the fact that the yearly ritual of evaluating (and sometimes rating and ranking) the performance of employees epitomizes the absurdities of corporate life. Managers and staff alike too often view performance management as time consuming, excessively subjective, demotivating, and ultimately unhelpful. In these cases, it does little to improve the performance of employees. It may even undermine their performance as they struggle with ratings, worry about compensation, and try to make sense of performance feedback.

These aren't new issues, but they have become increasingly blatant as jobs in many businesses have evolved over the past 15 years. More and more positions require employees with deeper expertise, more independent judgment, and better problem-solving skills. They are shouldering ever-greater responsibilities in their interactions with customers and business partners and creating value in ways that industrial-era performance-management systems struggle to identify. Soon enough, a ritual most executives say they dislike will be so outdated that it will resemble trying to conduct modern financial transactions with carrier pigeons.

Yet nearly nine out of ten companies around the world continue not only to generate performance scores for employees but also to use them as the basis for compensation decisions.[1] The problem that prevents managers' dissatisfaction with the process from actually changing it is uncertainty over what a revamped performance-management system ought to look like. If we jettison year-end evaluations—well, then what? Will employees just lean back? Will performance drop? And how will people be paid?

Answers are emerging. Companies, such as GE[2] and Microsoft,[3] that long epitomized the "stack and rank" approach have been blowing up their annual systems for rating and evaluating employees and are instead testing new ideas that give them continual feedback and coaching. Netflix[4] no longer measures its people against annual objectives, because its objectives have become more fluid and can change quite rapidly. Google transformed the way it compensates high performers at every level.[5] Some tech companies, such as Atlassian,[6] have automated many evaluation activities that managers elsewhere perform manually.

The changes these and other companies are making are new, varied, and, in some instances, experimental. But patterns are beginning to emerge.

• Some companies are rethinking what constitutes employee performance by focusing specifically on individuals who are a step function away from average—at either the high or low end of performance—rather than trying to differentiate among the bulk of employees in the middle.

• Many companies are also collecting more objective performance data through systems that automate real-time analyses.

• Performance data are used less and less as a crude instrument for setting compensation. Indeed, some companies are severing the link between evaluation and compensation, at least for the majority of the workforce, while linking them ever more comprehensively at the high and low ends of performance.

[1] See "The measure of a man," *Economist*, February 20, 2016.

[2] "Why GE had to kill its annual performance reviews after more than three decades," *Quartz*, August 13, 2015, qz.com.

[3] Nick Wingfield, "Microsoft abolishes employee evaluation system," *New York Times*, November 13, 2013, nytimes.com.

[4] Patty McCord, "How Netflix reinvented HR," *Harvard Business Review*, February 2014, hbr.org.

[5] Richard Feloni, "Inside Google's policy to 'pay unfairly'—why 2 people in the same role can earn dramatically different amounts," *Business Insider*, April 11, 2015, businessinsider.com.

[6] "8 automations that improved our HR team's productivity," *Atlassian blogs*, blog entry by jluijke, November 29, 2011, atlassian.com.

- Better data back up a shift in emphasis from backward-looking evaluations to fact-based performance and development discussions, which are becoming frequent and as-needed rather than annual events.

How these emerging patterns play out will vary, of course, from company to company. The pace of change will differ, too. Some companies may use multiple approaches to performance management, holding on to hardwired targets for sales teams, say, while shifting other functions or business units to new approaches.

But change they must.

RETHINKING PERFORMANCE

Most corporate performance-management systems don't work today, because they are rooted in models for specializing and continually optimizing discrete work tasks. These models date back more than a century, to Frederick W. Taylor.

Over the next 100 years, performance-management systems evolved but did not change fundamentally. A measure like the number of pins produced in a single day could become a more sophisticated one, such as a balanced scorecard of key performance indicators (KPIs) that link back to overarching company goals. What began as a simple mechanistic principle acquired layers of complexity over the decades as companies tried to adapt industrial-era performance systems to ever-larger organizations and more complicated work.

What was measured and weighted became ever more micro. Many companies struggle to monitor and measure a proliferation of individual employee KPIs—a development that has created two kinds of challenges. First, collecting accurate data for 15 to 20 individual indicators can be cumbersome and often generates inaccurate information. (In fact, many organizations ask employees to report these data themselves.) Second, a proliferation of indicators, often weighted by impact, produces immaterial KPIs and dilutes the focus of employees. We regularly encounter KPIs that account for less than 5 percent of an overall performance rating.

Nonetheless, managers attempt to rate their employees as best they can. The ratings are then calibrated against one another and, if necessary, adjusted by distribution guidelines that are typically bell curves (Gaussian distribution curves). These guidelines assume that the vast majority of employees cluster around the mean and meet expectations, while smaller numbers over- and underperform. This model typically manifests itself in three-, five-, or seven-

point rating scales, which are sometimes numbered and sometimes labelled: for instance, "meets expectations," "exceeds expectations," "far exceeds expectations," and so on. This logic appeals intuitively ("aren't the majority of people average by definition?") and helps companies distribute their compensation ("most people get average pay; overperformers get a bit more, underperformers a bit less").

But bell curves may not accurately reflect the reality. Research suggests that talent-performance profiles in many areas—such as business, sports, the arts, and academia—look more like power-law distributions. Sometimes referred to as Pareto curves, these patterns resemble a hockey stick on a graph. (They got their name from the work of Vilfredo Pareto, who more than a century ago observed, among other things, that 20 percent of the pods in his garden contained 80 percent of the peas.) One 2012 study concluded that the top 5 percent of workers in most companies outperform average ones by 400 percent. (Industries characterized by high manual labor and low technology use are exceptions to the rule.[7]) The sample curve emerging from this research would suggest that 10 to 20 percent of employees, at most, make an outsized contribution.

Google has said that this research, in part, lies behind a lot of its talent practices and its decision to pay outsized rewards to retain top performers: compensation for two people doing the same work can vary by as much as 500 percent.[8] Google wants to keep its top employees from defecting and believes that compensation can be a "lock-in"; star performers at junior levels of the company can make more than average ones at senior levels. Identifying and nurturing truly distinctive people is a key priority given their disproportionate impact.

Companies weighing the risks and rewards of paying unevenly in this way should bear in mind the bigger news about power-law distributions: what they mean for the great majority of employees. For those who meet expectations but are not exceptional, attempts to determine who is a shade better or worse yield meaningless information for managers and do little to improve performance. Getting rid of ratings—which demotivate and

[7] Ernest O'Boyle Jr. and Herman Aguinis, "The best and the rest: Revisiting the norm of normality of individual performance," *Personal Psychology*, 2012, 65, pp. 79–119. Researchers canvassed studies involving more than 600,000 people in academia, politics, entertainment, and sports. They found performance power curves consistent across different jobs, performance measures, and time frames.

[8] Google's senior vice president for people operations, Laszlo Bock, wrote about these practices in his book, *Work Rules: Insights from Inside Google That Will Transform How You Live and Lead*, New York, NY: Hachette Book Group, 2015.

irritate employees, as researchers Bob Sutton and Jeff Pfeiffer have shown—makes sense.

Many companies, such as GE, the Gap,[9] and Adobe Systems,[10] have done just that in a bid to improve performance. They've dropped ratings, rankings, and annual reviews, practices that GE, for one, had developed into a fine art in previous decades. What these companies want to build—objectives that are more fluid and changeable than annual goals, frequent feedback discussions rather than annual or semiannual ones, forward-looking coaching for development rather than backward-focused rating and ranking, a greater emphasis on teams than on individuals—looks like the exact opposite of what they are abandoning.

The point is that such companies now think it's a fool's errand to identify and quantify shades of differential performance among the majority of employees, who do a good job but are not among the few stars. Identifying clear overperformers and underperformers is important, but conducting annual ratings rituals based on the bell curve will not develop the workforce overall. Instead, by getting rid of bureaucratic annual-review processes—and the behavior related to them—companies can focus on getting much higher levels of performance out of *many more* of their employees.

GETTING DATA THAT MATTER

Good data are crucial to the new processes, not least because so many employees think that the current evaluation processes are full of subjectivity. Rather than relying on a once-a-year, inexact analysis of individuals, companies can get better information by using systems that crowdsource and collect data on the performance of people and teams. Continually crowd-sourcing performance data throughout the year yields even better insights.

For instance, Zalando, a leading European e-retailer, is currently implementing a real-time tool that crowdsources both structured and unstructured performance feedback from meetings, problem-solving sessions, completed projects, launches, and campaigns. Employees can request feedback from supervisors, colleagues, and internal "customers" through a real-time online app that lets people provide both positive and more critical comments about each other in a playful and engaging way. The system then weights responses by how much exposure the provider has to the requestor. For every kind of behavior that employees seek or provide

9 Vauhini Vara, "The push against performance reviews," *New Yorker*, July 24, 2015, newyorker.com.

10 *Adobe Life Blog*, "The dreaded performance review? Not at Adobe," adobe.com.

feedback about, the system—a structured, easy-to-use tool—prompts a list of questions that can be answered intuitively by moving a slider on the touchscreen of a mobile device. Because the data are collected in real time, they can be more accurate than annual reviews, when colleagues and supervisors must strain to remember details about the people they evaluate.

Employees at GE now use a similar tool, called PD@GE, which helps them and their managers to keep track of the company's performance objectives even as they shift throughout the year. The tool facilitates requests for feedback and keeps a record of when it is received. (GE is also changing the language of feedback to emphasize coaching and development rather than criticism.) GE employees get both quantitative and qualitative information about their performance, so they can readjust rapidly throughout the year. Crucially, the technology does not replace performance conversations between managers and employees. Instead, these conversations center around the observations of peers, managers, and the employees themselves about what did and didn't help to deliver results. GE hopes to move most of its employees to this new system by the end of 2016.

In other words, tools can automate activities not just to free up time that managers and employees now spend inefficiently gathering information on performance but also transform what feedback is meant to achieve. The quality of the data improve, too. Because they are collected in real time from fresh performance events, employees find the information more credible, while managers can draw on real-world evidence for more meaningful coaching dialogues. As companies automate activities and add machine learning and artificial intelligence to the mix, the quality of the data will improve exponentially, and they will be collected much more efficiently.[11]

Finally, performance-development tools can also identify the top performers more accurately, though everyone already knows subjectively who they are. At the end of the year, Zalando's tool will automatically propose the top 10 percent by analyzing the aggregated feedback data. Managers could adjust the size of the pool of top performers to capture, say, the best 8 or 12 percent of employees. The tool will calculate the "cliff" where performance is a step function away from that of the rest of the population. Managers will therefore have a fact-based, objective way to identify truly distinctive employees. Companies can also use such systems to identify those who have genuinely fallen behind.

[11] For additional insights, see Aaron De Smet, Susan Lund, and William Schaninger, "Organizing for the future," *McKinsey Quarterly*, January 2016, McKinsey.com.

Relatively easy and inexpensive to build (or to buy and customize), such performance-development applications are promising—but challenging (see the exhibit for a generic illustration of such an app). Employees could attempt to game systems to land a spot among the top 10 percent or to ensure that a rival does not. (Artificial intelligence and semantic analysis might conceivably distinguish genuine from manicured performance feedback, and raters could be compared with others to detect cheating.) Some employees may also feel that Big Brother is watching (and evaluating) their every move. These and other real-life challenges must be addressed as more and more companies adopt such tools.
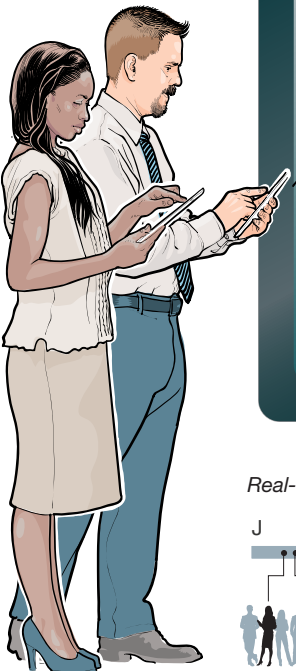
TAKE THE ANXIETY OUT OF COMPENSATION

The next step companies can take to move performance management from the industrial to the digital era is to take the anxiety out of compensation. But this move requires managers to make some counterintuitive decisions.

Exhibit

Continually crowdsourcing performance data provides fresher and more timely insights.

An online app facilitates the collection of real-time peer feedback from multiple sources throughout the year.

Q Kneesund

Hans Kneesund
What did you think of my presentation?

RUBBISH   GREAT

Award Hans a badge for ...

Speedy Brainy Listens Team player

FEEDBACK:

Great use of infographics in that presentation, Hans.
LISA LOTTABEE

- Structure feedback along set performance dimensions or treat more informally.

- Request feedback at any time—eg, from leader, team member, or customer.

- Use badges and comment field for additional nuances.

- Offer unprompted feedback.

Real-time feedback, throughout the year, from multiple sources

J F M A M J J A S O N D
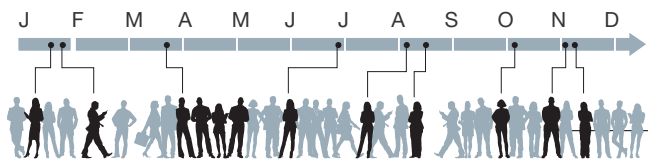
Conventional wisdom links performance evaluations, ratings, and compensation. This seems completely appropriate: most people think that stronger performance deserves more pay, weaker performance less. To meet these expectations, mean performance levels would be pegged around the market average. Overperformance would beat the market rate, to attract and retain top talent. And poor scores would bring employees below the market average, to provide a disincentive for underperformance. This logic is appealing and consistent with the Gaussian view. In fact, the distribution guide, with its target percentages across different ratings, gives companies a simple template for calculating differentiated pay while helping them to stay within an overall compensation budget. No doubt, this is one of the reasons for the prevalence of the Gaussian view.

This approach, however, has a number of problems. First, the cart sometimes goes before the horse: managers use desired compensation distributions to reverse engineer ratings. To pay Tom x and Maggie y, the evaluator must find that Tom exceeds expectations that Maggie merely meets. That kind of reverse engineering of ratings from *a priori* pay decisions often plays out over several performance cycles and can lead to cynical outcomes—"last year, I looked out for you; this year, Maggie, you will have to take a hit for the team." These practices, more than flaws in the Gaussian concept itself, discredit the performance system and often drown out valuable feedback. They breed cynicism, demotivate employees, and can make them combative, not collaborative.

Second, linking performance ratings and compensation in this way ignores recent findings in the cognitive sciences and behavioral economics. The research of Nobel laureate Daniel Kahneman and others suggests that employees may worry excessively about the pay implications of even small differences in ratings, so that the fear of potential losses, however small, should influence behavior twice as much as potential gains do. Although this idea is counterintuitive, linking performance with pay can demotivate employees even if the link produces only small net variances in compensation.

Since only a few employees are standouts, it makes little sense to risk demotivating the broad majority by linking pay and performance. More and more technology companies, for instance, have done away with performance-related bonuses. Instead, they offer a competitive base salary and peg bonuses (sometimes paid in shares or share options) to the company's overall performance. Employees are free to focus on doing great

work, to develop, and even to make mistakes—without having to worry about the implications of marginal rating differences on their compensation. However, most of these companies pay out special rewards, including discretionary pay, to truly outstanding performers: "10x coders get 10x pay" is the common way this principle is framed. Still, companies can remove a major driver of anxiety for the broad majority of employees.

Finally, researchers such as Dan Pink say that the things which *really* motivate people to perform well are feelings like autonomy, mastery, and purpose. In our experience, these increase as workers gain access to assets, priority projects, and customers and receive displays of loyalty and recognition. Snapping the link between performance and compensation allows companies to worry less about tracking, rating, and their consequences and more about building capabilities and inspiring employees to stretch their skills and aptitudes.

A large Middle Eastern technology company recently conducted a thorough study of what motivates its employees, looking at combinations of more than 100 variables to understand what fired up the best people. Variables studied included multiple kinds of compensation, where employees worked, the size of teams, tenure, and performance ratings from colleagues and managers. The company found that *meaning*—seeing purpose and value in work—was the single most important factor, accounting for 50 percent of all movement in the motivation score. It wasn't compensation. In some cases, higher-paid staff were markedly less motivated than others. The company halted a plan to boost compensation by $100 million to match its competitors.

Leaders shouldn't, however, delude themselves into thinking that cutting costs is another reason for decoupling compensation from performance evaluations. Many of the companies that have moved in this direction use generous stock awards that make employees up and down the line feel not only well compensated but also like owners. Companies lacking shares as currency may find it harder to make the numbers work unless they can materially boost corporate performance.

COACHING AT SCALE TO GET THE BEST FROM THE MOST

The growing need for companies to inspire and motivate performance makes it critical to innovate in coaching—and to do so at scale. Without great and frequent coaching, it's difficult to set goals flexibly and often, to help employees stretch their jobs, or to give people greater responsibility and autonomy while demanding more expertise and judgment from them.

Many companies and experts are exploring how to improve coaching—a topic of the moment. Experts say three practices that appear to deliver results are to change the language of feedback (as GE is doing); to provide constant, crowdsourced vignettes of what worked and what didn't (as GE and Zalando are); and to focus performance discussions more on what's needed for the future than what happened in the past. Concrete vignettes, made available just in time by handy tools—and a shared vocabulary for feedback—provide a helpful scaffolding. But managers unquestionably face a long learning curve for effective coaching as work continues to change and automation and reengineering configure job positions and work flows in new ways.

———————

Companies in high-performing sectors, such as technology, finance, and media, are ahead of the curve in adapting to the future of digital work. So it's no surprise that organizations in these sectors are pioneering the transformation of performance management. More companies will need to follow—quickly. They ought to shed old models of calibrated employee ratings based on normal distributions and liberate large parts of the workforce to focus on drivers of motivation stronger than incremental changes in pay. Meanwhile, companies still have to keep a keen eye on employees who are truly outstanding and on those who struggle.

It's time to explore tools to crowdsource a rich fact base of performance observations. Ironically, companies like GE are using technology to democratize and rehumanize processes that have become mechanistic and bureaucratic. Others must follow. ⓠ

Boris Ewenstein, Bryan Hancock, and **Asmus Komm** are expert principals in McKinsey's Johannesburg, Atlanta, and Hamburg offices, respectively.