

# Culture . Rocks

## PERFORMANCE MANAGEMENT at GOOGLE



# **Case Study: How Google does Performance Reviews**

Everything you need to know about Google's performance management practices

Francisco S. homem de Mello

© 2016 Francisco S. homem de Mello

# Contents

- Introduction . . . . . 1
- Where did we take all this stuff from? . . . . . 3
- Overview . . . . . 4
- Performance Reviews . . . . . 5
  - Self-Evaluation . . . . . 6
  - 360-degree Feedback . . . . . 6
  - Calibration . . . . . 7
  - Outputs . . . . . 8
- OKRs . . . . . 12
  - TL;DR . . . . . 12
  - Introduction . . . . . 13
  - A Brief History of OKRs . . . . . 14
  - A bit of goal-setting science . . . . . 17

# Introduction

*"We need people to know how they're doing, and we've evolved what might at first seem like a zany complex system that shows them where they stand. Along the way, we learned some startling stuff. We're still working on it, as you'll see, but I feel pretty confident we're headed in the right direction. And with any luck I can save you some of the headaches and missteps we had along the way."*

- Laszlo Bock, SVP, People Operations, Google

Google has probably Silicon Valley's, and maybe the world's, most advanced human resources (or, as they call it, People Operations) practice. As it becomes clear by books like *Work Rules*, written by its SVP of People Operations, and *How Google Works*, by Eric Schmidt, its former CEO, and Jonathan Rosenberg, current SVP of Product, the company continuously iterates on people practices, based on uniquely huge amounts of data, gathered among its more than 50 thousand "smart creatives", employees in the fields of engineering, design, and sales, all handpicked at the world's top universities.

Google's people operations cornerstones are:

- Hiring only the best: sourcing, and selecting, only the best fit candidates amongst the best pool of candidates worldwide, and, if it can't reasonably achieve 100% perfection in hiring only amazing fits, skewing errors towards false negatives (eventually passing on a great candidate) instead of false positives (eventually hiring a bad fit);
- Creating a meritocratic environment, where the best performances are correctly identified and rewarded; and,

- Developing employees to their full potential, through great people management and on-the-job coaching (see our Project Oxygen paper here), peer-to-peer and outside training, and through a comprehensive 360-degree feedback collection process;

The purpose of this case is to explain a bit better, and in detail, how Google does the second bullet point, meritocracy, and bits of the third, development, through its performance management procedures.

But why, you may ask, look at Google for a benchmark? Apart from the obvious reasons (it seems to be working for them, eh?), it our view that smaller, less resourceful companies (and here we're talking more about sheer cash and headcount, as opposed to attitude), can greatly benefit from using Google as a starting point for their own practices, and then iterating on that (what is commonly known as standing on the shoulders of giants). Why should you, HR manager, or C-level executive, reinvent the wheel when this giant company has not only spent millions and millions of dollars finding its best self, but talked at length about it, so that you can benchmark yourself and use many of these practices? There're no good reasons not to[1].

We'll have achieved our goals if you find some inspiration and best practices on this paper. Remember we, at Culture.Rocks, can help you get your performance management practices (be them inspired by Google or not) running in a matter of hours.

# **Where did we take all this stuff from?**

Great question! Google, as many of the world's top companies (like GE, AB InBev, Walmart, etc), talks frequently and openly about its own culture. So we basically read everything that's out there – and actually written by Google and its current and former employees -. We've also scanned online platforms like Quora, Medium, and Twitter, read many pieces by the press, and, finally, interviewed as many as 10 former executives, both from People Operations and from areas like Search and Google Ventures, in order to form a holistic understanding of Google's practices.

We're confident that you have a very faithful description in your hands. Some of the processes may have already been iterated out activity, but for the most part, we strongly believe you have an accurate picture of the company's current practices.

# Overview

For the purposes of this case study, we're calling performance management the collection of the following human resources tools and processes used at Google:

- Annual performance review (including mid-year checkpoint)
- Monthly performance check-ins (part of regular 1:1 meetings that also comprise other themes such as career development, coaching, personal issues, etc.)
- Googlegeist engagement survey (that spans much more than just the regular engagement axes, but measures basically everything that's to be measured)
- Annual Upward Feedback Survey, a feedback review (similar to 360-degree review) where only supervisors are reviewed by their direct reports, and that is based on Google's Project Oxygen
- OKRs, or objectives and key-results, a mildly different form of Management-by-Objectives, that we explain in this post, and
- Meritocracy, or compensating people unequally, based on their perceived performance, through bonuses, equity stock-option grants, and prizes Performance reviews

# Performance Reviews

Google's annual performance review cycle is comprised of two parts: a "preview", in the end of the first semester, and a complete review, that happens between October and November, and which happens concurrently with the company's 360-degree feedback collection process.

Managers take two main things into account when attributing their employees' performance ratings: results attained, or what the employee accomplished, and behaviors, or how the employee attained these results. The employee starts with a self-assessment, which is followed by peer-reviews, whose authors are only visible to managers (reviewees may have access to the anonymized content of peer reviews).

On the review side, Google employees are asked to review each other, and their direct reports, according to the following criteria:

- **Googleyness:** The employee's adherence to Google's values. This is the main component of the "how" axis.
- **Problem solving:** Analytical skills applied to work situations (problem solving).
- **Execution (high quality work with little guidance):** Delivering great work without the need for a lot of hand-holding from managers and peers (autonomy).
- **Thought leadership:** How much an employee is seen as a reference for a given niche of expertise. As Google grows in size, these niches may tend to become smaller and smaller, but still, Google wants employees that are go-to resources for specific themes, training colleagues on tech-talks, training customers, and producing high-quality content.



- **Leadership (or emerging leadership):** Albeit many young Googlers have little or no exposure to managing complex teams, everybody is required, nonetheless, to show emerging leadership skills, such as taking the lead of problems and projects, being pro-active, and owning results personally.
- **Presence:** Presence is the employee's ability to make himself heard in an increasingly large organization, and intimately related to emerging leadership.

## Self-Evaluation

The self-evaluation is the first step in the performance review, and where the employee evaluates himself in the five criteria described above (on five grades ranging from “never demonstrates” all the way to “always demonstrates” and invited to share examples of his actions that support these grades), and highlights his main accomplishments for the last cycle (in a text field limited to 512 characters). These accomplishments will appear in the next step (360-degree reviews) to reviewing peers, who'll be then asked to assess their proximity with these projects, and the reviewee's impact on their results.

## 360-degree Feedback

Google's 360-degree review process serves the purpose of giving managers a holistic picture of their direct reports, since they may carry a biased and restricted impression of reports' impact and behavior (some employees may be great at “managing up” a rosy picture of their contributions, for example).

The process starts with a back-and-forth between employee and manager, so as to pick a representative, fair sample of peers to participate. The employee suggests a shortlist, that is discussed and

validated with the manager, taking into account how close the peer was to the employee's contributions, and how well she can assess the employee's performance.

Peers are expected to give assessments in three different media: strengths, or things that the person should keep on doing, and weaknesses, or things that the person should consider working on/developing; rating each other on the five criteria discussed above; and finally, commenting on the reviewee's contribution to specific projects. These two open-ended fields (positives and negatives) have evolved from a larger form a few years ago. Laszlo Bock, Google's SVP, People Operations, observes in his Work Rules that the simplification reduced aggregate time spent on this step by more than 25%, while improving the share of participants who perceived it as useful from 49% to 75%.

## Calibration

After all data has been collected, in the form of self-reviews and peer-reviews (or what's known as 360-degree feedback), and results achieved are understood, managers draft a rating for their employees, based on the following scale[2]:

- Needs improvement
- Consistently meets expectations
- Exceeds expectations
- Strongly exceeds expectations
- Superb

As you may have noticed, I said they draft their ratings. That's because no ratings are final before the calibration process, again, described by Laszlo Bock:

*"The soul of performance assessment is calibration... A manager assigns a draft rating to an employee – say, 'exceeds expectations' –*

*based on mainly OKRs but tempered by other activities, like the volume of interviews completed, or extenuating circumstances such as a shift in the economy that might have affected ad revenues. Before his draft rating becomes final, groups of managers sit down together and review all of their employees' draft ratings together in a process we call calibration... A group of five to ten managers meet and project on a wall their fifty to a Thousand employees, discuss individuals, and agree on a fair rating. This allows us to remove the pressure managers may feel from employees to inflate ratings. It also ensures that the end results reflect a shared expectation of performance, since managers often have different expectations for their people and interpret performance standards in their own idiosyncratic manner... Calibration diminishes bias by forcing managers to justify their decisions to one another. It also increases perceptions of fairness among employees."*

Calibration, a process also adopted at other leading companies such as AB InBev, GE, Kraft Heinz, and Goldman Sachs, is therefore of crucial importance in ensuring the fairness of performance ratings. It's where heavy-handed raters are identified and discounted for (and the opposite is also true).

## Outputs

The calibration meetings output each and every employee's performance rating for the period. After the rating is closed, managers go on to hold two meetings: one where feedback is given, taking into account peer reviews and managers' impressions of their employees, and another where compensation and promotion decisions are communicated.

The two conversations are held in different meetings and at least a month apart from each other in order to ensure their quality. Google understands that a compensation-focused employee is no good a listener of feedback, whether compensation expectations were not

met, met, or exceeded:

“a [negative] dynamics exists when managers sit down to give employees their annual review and salary increase. The employees focus on the extrinsic reward – a raise, higher rating – and learning shuts down.... We have an embarrassingly simple solution. Never have the [pay and feedback] conversations at the same time. Annual reviews happen in November, and pay discussions happen a month later.”

The theme is also discussed by Prasad Setty, member of Google’s People & Innovation Lab[3]:

*“Traditional performance management systems make a big mistake. They combine two things that should be completely separate: performance evaluation and people development. Evaluation is necessary to distribute finite resources, like salary increases or bonus dollars. Development is just as necessary for so people grow and improve.”*

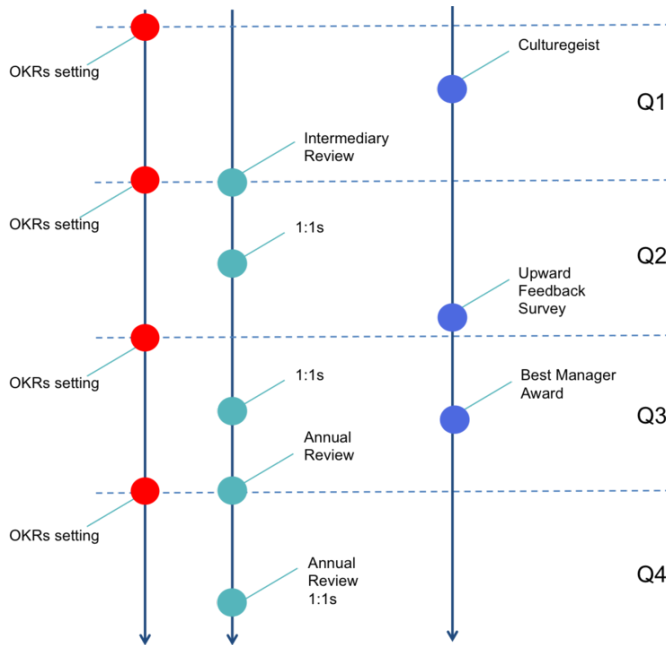


Diagram 1: Google's simplified performance management schedule

[1] Actually there is one good reason not to: You're in a business where the majority of your employees are not "smart creatives," but maybe less educated, operational, hourly workers, maybe not as capable of self-management, and maybe not as high on Maslow's pyramid. Valid argument, but we won't discuss it in detail here. Enough to say that you'll have much more to gain from learning with Google than ignoring it, for now.

[2] Before a five-point scale, Google rated its employees on a scale from 1 to 5 in 0,1 increments, having, in fact, 40+ possible ratings. The scale, according to Laszlo Bock, bore many inefficiencies, as was ditched after more than 10 years in use for a simpler scale.

[3] Google's People & Innovation Lab, or PiLab, is worth a book itself. In short, it's a team of quants whose only attribution is to study people data (performance, engagement, happiness, etc),

iterate on people practices (testing them), and to continuously support Google's people practices with heavyweight data analytics.

# OKRs

\*This chapter is an excerpt from our ebook, The Ultimate Guide to OKRs, which you can find and download at <http://qulture.rocks/the-ultimate-guide-to-okrs>

## TL;DR

OKRs are an acronym for Objectives and Key Results. Objectives are high-level, qualitative goals. Key-Results are specific, SMART goals that support the Objective. When we say support, we mean Key-Results should include metrics that trully translate Objective accomplishment.

Some pundits use a very simple statement: We will achieve \_\_\_\_\_ - \_\_\_\_\_ as measured by \_\_\_\_\_, \_\_\_\_\_, *and* \_\_\_\_\_. The first space is filled by your Objective, and the second to fourth are filled by Key-Results. Let's use an example to illustrate our definition:

### Objective

- Increase the profitability of the company

(Since OKRs belong to cycles, if they don't have a "date" stamp to them, you should automatically assume the goals should be completed before the end of the cycle.)

### Key Results

- Increase revenues by 10%
- Reduce costs by 3%

- Maintain general, and administrative expenses nominally constant

As you can see, the Objective is a bold goal, specific, time bound, but still achievable (as opposed to doubling my profit in a month).

Key results are the actual targets, as measured by KPIs, that will really acid test Objective achievement. They are what, in MBOs, we know by “goals”. As you can see, I’ve used Andy Grove’s method of Key-Results breakdown: using KRs as milestones for goal achievement. There are other possible approaches. Some people defend that Objectives have to be qualitative, whereas Key-Results have to be quantitative.

## Introduction

If we come to think about it, there’s a blurred line that separates a job description, a goal, and standard operating procedures.

Job descriptions would be fine by themselves if we substituted humans with robots, and robots were always pre-programmed with job descriptions and a manual of standard operating procedures: they would perform their jobs and responsibilities perfectly, capped only by the machines mechanical/processing limitations. But it turns out there are many tasks we haven’t figured how to program robots to do better than us, yet. We humans are very good at having common sense, and transferring knowledge from one set of applications to another; at figuring out weird patterns; and at coming up with novel solutions to problems.

With this set of amazing skills comes also some limitations at performing and sustaining performance at our fullest potential, two problems that are closely related to our - unique? - faculty: volition. That could happen because a lack of challenge (related to our frontal cortex) or a lack of purpose - the “why” of what we’re doing it. If



we're supplied only with job descriptions, we may forget them, or lose our motivation to perform it at our best. We may even lose track of what "best" means.

Goals come in to fill this need. They help us find a sense of purpose when we properly participate in their setting, but also by linking our responsibilities at tasks' outcomes to greater outcomes, like the success of the group we work in/with; they provide us with a challenge to constantly perform at our best; they provide gratification when we reach them, which reinforces the cycle.

OKRs are basically goals: an old staple of business management, rebranded, repurposed, and tweaked to 21st century necessities of companies and professionals.

## **A Brief History of OKRs**

OKRs are an old staple of business management, rebranded, repurposed, and tweaked to 21st century necessities of companies and professionals.

It all started with the fathers of management, Taylor, Ford, and the sorts, who began facing business like a science. How so? They figured they could measure outcomes, and then formulate hypotheses as to how they could improve these outcomes. The main outcomes back then were productivity, as measured by output per employee. These guys figured out optimum work schedules, break times, and lightning arrangements for factories. They also started streamlining production, adding specialization to the factory floor. These practices all brought incredible, tangible improvements.

In the 50s, a fellow named Peter Drucker, who's believed to be the greatest management guru that ever lived, figured out that adding goals to managers could be a great thing. Not only they had to improve their outcomes, as Taylor and Ford had, but they also had to aim at specific target outcomes from time to time. Drucker called

this framework Management by Objectives, or MBO, a concept introduced in his seminal book *The Practice of Management*.

Since the introduction of MBOs, practically every modern Fortune 500 company practices some sort of goal setting. It's proven to bring better results than not having goals. Some companies set them once a year; some companies set them twice a year. A number of them tie variable compensation to reaching your goals, and another number of them perform some sort of performance review based on these goals and results. The term OKRs was introduced by Andy Grove, a de facto cofounder of Intel (he joined the company on the day of its incorporation, but is not listed as a cofounder,) and its former CEO, in his great management book *High Output Management*. Grove didn't bring any transformational insight to MBOs, but spoke about appending key-results to goals, and calling goals objectives. But making key-results an integral part of the MBO process is very important: it brings clarity to how goals can and should be attained, and makes this "how" evident and transparent to everybody.

In Grove's view, key-results had to be chronological milestones that took professionals in the direction of reaching their goals: a one-year goal could be broken down into 12 monthly key-results, or 4 quarterly key-results. He treated them specifically as "milestones." That use was attuned to Intel's 80s reality: a large company, already on the top of its game, looking to translate its strategic planning into actionable goals and milestones for the whole organization.

Another Grove tweak to MBOs was his belief that goals (objectives) and key-results had to be set in a bottom-up process, from the employee up, so as to bring buy-in and empowerment to the process. Before that, companies would shove goals down the organization, from the Board to the CEO, down to VPs, and so on. Grove enabled employees to set their goals according to a broad guidance from the company, to be then calibrated with direct managers.

Last, but not least, Grove insisted that OKRs be aggressive, meaning – very – hard to achieve, what he called "stretched." He went further

along, and instituted 70% as the new 100%, meaning that achieving 70% of your goals was as good as hitting them, since they were purposely baked very hard .

In the late 90s, OKRs spread out to other Silicon Valley companies, through the inspiration of Jon Doerr, a partner of Kleiner Perkins Caufield Byers, one of the world's foremost venture capital firms. Doerr had worked for Intel under Grove's leadership, and got acquainted with its use of OKRs, later thinking it could be adapted to other companies of KP's portfolio. That's how Google, and later on Zynga, became fierce advocates of OKRs, tweaking the tool to their specific needs.

But what made Google's version of OKRs different from Intel's? Not much. Google shortened – a lot – the OKR cycle, making it a quarterly process. That means the company, its senior executives, and basically every employee, sets his or hers objectives and corresponding key-results quarterly, a practice more attuned to the incredibly fast-paced reality of web 2.0 technology companies. Google enforced Grove's position that goals should not be cascaded down the organization in a top-down manner, and greatly expanded upon it, according to Laszlo Bock, its SVP of People Operations:

*“Having goals improves performance. Spending hours cascading goals up and down the organization, however, does not. It takes way too much time and it's too hard to make sure all the goals line up. We have a market-based approach, where over time our goals converge, because the top OKRs are known and everyone else's OKRs are visible. Teams that are grossly out of alignment stand out, and the few major initiatives that touch everyone are easy enough to manage directly.”*

That means at Google, everyone's OKRs are set by themselves, and made public via its intranet. Google ensures that individual OKRs are aligned with its own through a mix of supervisor oversight, peer pressure, and psychology.

## A bit of goal-setting science

(Or, the goal of having goals.)

Goal-setting has historically been used in the corporate world for two main purposes:

- To motivate employees (efficiency)
- To assess their performance

Let me explain this: HR common sense has always said that goals motivate employees towards achieving better results. Goal achievement, on the other hand, has historically been used as a proxy for performance: if I've hit 100% of my goals, it must mean I'm a good performer. But we thought it made sense to briefly review goal-setting theory, or GST. We think HR professionals deserve to have this widespread practice correctly understood with a theoretical basis, because there is more to it than just these two axes of purpose.

According to GST, <sup>1</sup>foo3 goals serve three main purposes:

### Focus

Presuming that goals have been come up with according to the company's long, medium, and short-term strategies, according to myriad methodologies like BSC, Hoshin Kanri, etc, goals help the company focus effort, attention, and energy on what's relevant, relative to what's not relevant. According to Johnson, Chang, and Lord (2006), "goals direct individuals' attention to goal-relevant activities and away from goal-irrelevant activities." It is proven that "individuals cognitively and behaviorally pay more attention to a task that is associated with a goal than to a task that is not."

---

<sup>1</sup>MarionEberly,DongLiu,TerenceMitchell,ThomasLee;  
AttributionsandEmotionsasMediatorsand/orModeratorsintheGoal-StrivingProcess

## **Effort**

Another very important purpose of goals is to increase the level of effort that people exert at work. It is also proven that “goals energise and generate effort toward goal accomplishment. The higher the goal, the more the effort exerted.” This is a tricky equation: too hard a goal, and, as you’ll see in a bit, employees get demotivated; too easy a goal, and employees will also get demotivated. In sum, there’s a right amount of hard, which pushes people to challenge themselves, but within a reasonable chance of achievement, that optimizes performance, which links us to

## **Persistence**

Persistence is probably the trickiest thing to get right when setting goals: The right ones produce high effort input for longer periods of time, but the wrong ones can really wreak havoc: “large negative discrepancies may lead to a withdrawal of effort when individuals are discouraged and perceive low likelihood of future goal attainment.” (Carver & Scheier, 1998). As we’ll see, there are derivative factors that influence persistence towards goals.