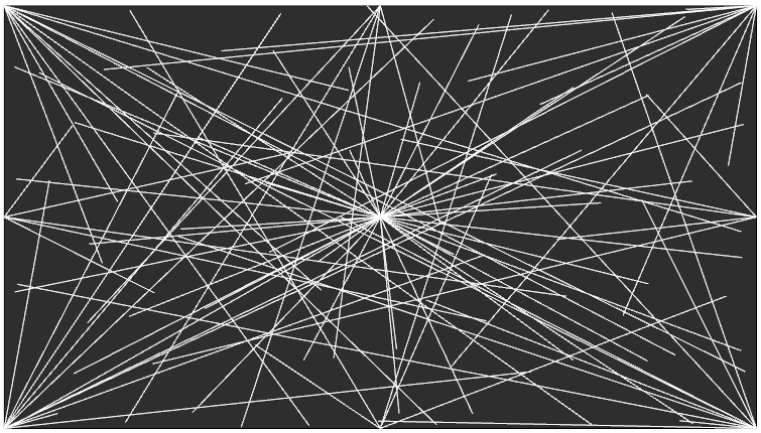
Why We Invest in Reviews and Rate Performance

Published on Dec 07, 2016 by Niko Canner



GE made headlines with their decision to eliminate ratings.

Accenture's CEO announced in an interview with the Washington Post last year that they were eliminating their annual reviews: "We're going to get rid of it. Not 100 percent, but we're going to get rid of probably 90 percent of what we did in the past. It's not what we need. We are not sure that spending all that time on performance management has been yielding a great outcome."

Progressive sentiment in the corporate world is increasingly that, to quote GE CHRO Sue Peters, "the world isn't really on an annual cycle anymore for anything." Performance reviews are called a relic of an earlier era, with an impact on actual company performance that rates somewhere between wasteful and harmful.

So why on earth are we at Incandescent, with no legacy systems to preserve, passion for innovation in management systems, and a team largely made up of Millennials, spending significant time and energy writing annual reviews that include ratings?

Understanding the Arguments Against Reviews

Before laying out the specifics of our decision to invest in reviews, let's spend a few minutes looking at the arguments against them. As Albert Hirschman brilliantly laid out in his book *The Rhetoric of Reaction*, arguments against a solution that's advanced as a way to make some positive impact – whether the French Revolution or the practice of performance reviews — generally take one of three forms (the below examples are provocatively illustrative and not representative of my own political positions):

- **Perversity**: the actual impact of the solution is to make worse the very thing that the solution aims to make better (e.g., the argument that raising the minimum wage will make unskilled

workers worse off by raising their rate of unemployment)
- **Futility**: the solution won't address an intractable problem, and will consume resources without positive effect (e.g., the argument that we can't make schooling better by regulating it at a national level, and therefore shouldn't try)
- **Jeopardy**: the solution will compromise something else we fundamentally value or should value (e.g., the argument that welfare programs will undermine the will to work and therefore the moral backbone of the recipients)

All three of these forms of argument show up in the phalanx arrayed against performance reviews over the past few years.

My friend David Rock is among the highest-profile voices of the perversity argument. Take for example this argument from his Strategy & Business piece "Kill Your Performance Ratings":

> *… labeling people with any form of numerical rating or ranking automatically generates an overwhelming "fight or flight" response…impairs good judgment. This neural response is the same type of "brain hijack" that occurs when there is an imminent physical threat like a confrontation with a wild animal. It primes people for rapid reaction and aggressive movement. But it is ill-suited for the kind of thoughtful, reflective conversation that allows people to learn from a performance review.*

For example, a supervisor might say, with the best of intentions, "You were ranked number 2 this year, and here are some development actions for the future." In this company, which scores its appraisals on a 1–3 scale, a 2 ranking is supposed to represent high praise. But a typical employee immediately disengages. Knowing that others were ranked still higher is enough to provoke a brain hijack. The employee may not say anything overtly, but he or she feels disregarded and undermined—and thus intensely inclined to ignore feedback, push back against stretch goals, and reject the example of positive role models.

Many of the voices from within the corporate world articulating the case against performance reviews make the futility argument. For instance, take this passage from near the end of Marcus Buckingham and Ashley Goodall's Harvard Business Review piece about Deloitte's reexamination of its performance management approach, "Reinventing Performance Management":

Over the past few years the debate about performance management has been characterized as a debate about ratings—whether or not they are fair, and whether or not they achieve their stated objectives. But perhaps the issue is different: not so much that ratings fail to convey what the organization knows about each person but that as presented, that knowledge is sadly one-dimensional. In the end, it's not the particular number we assign to a person that's the problem; rather, it's the fact that there *is* a single number. Ratings are a distillation of the truth—and up until now, one might argue, a necessary one. Yet we want our organizations to know us, and we want to know ourselves at work, and that can't be compressed into a single number.

To restate this in a way that emphasizes the theme of futility: the problem about ratings is that they don't actually achieve what they're meant to achieve, and in fact nothing that's like a "sadly one-dimensional" rating could achieve the goal.

The jeopardy argument moves from a focus on the empirical effects of a practice to a focus on more fundamental values. Take for example this passage from Ron Lebow and Randy Spitzer's book Accountability, quoted by Frederic Laloux in *Reinventing Organizations*:

Too often, appraisal destroys human spirit and, in the span of a 30-minute meeting, can transform a vibrant, highly committed employee into a demoralized, indifferent wallflower who reads the want ads on the weekend.... They don't work because most performance appraisal systems are a form of judgment and control.

The argument here isn't so much that the tactic doesn't work – although Lebow and Spitzer do assert perversity – as that something deeper is being compromised, that humanity is undermined by practices that embody "judgment and control."

Against the backdrop of these diverse arguments, our basic beliefs are:

1. In our culture, valuing our goals means that we value being as accurate as we can be in our understanding of what drives performance. This includes (but is not limited to) our understanding of what individuals contribute; where individuals contribute most and fall short; and what individuals can do to improve as rapidly as possible
2. In the context of paying for performance – people who contribute more and demonstrate more valuable capabilities get paid more – we owe everyone a transparent articulation of the assessments that impact their pay
3. Being explicit about assessments enables dialogue and action. Where there is disagreement, isolating and understanding that disagreement is essential to getting to better answers
4. Where there are negative consequences from discussions of performance – e.g., defensive reactions like the ones David Rock calls out – those can be processed and addressed in ways that strengthen everyone

To the "jeopardy" argument, we assert precisely the opposite: transparent judgments about performance are a natural part of a human, vital culture that celebrates shared commitment to both a larger purpose and the flourishing of people each on a quest to "become a better instrument" in service of their own deepest goals. In the face of the "futility" argument, we observe, empirically, that people take away valuable insight from a carefully designed approach to periodically stepping back and looking at a synthesized, "official" view of their performance and development. And we believe vis-à-vis the perversity argument that negative effects can easily be managed (e.g., defensive reactions are observed, understood, and worked through in a way that both creates openness to important information and strengthens the underlying "muscles" involved in avoiding defensiveness), and that the perverse effects of systems that withhold a clear synthesis of performance and how that synthesis relates to pay are far greater, at least in companies like ours where compensation is in fact differentiated based on performance.

The Devil is in the Details

We don't disagree with the assessment that most annual review systems work poorly, and certainly the data is clear that most employees in most companies experience these systems as dysfunctional.

We take these concrete steps to avoid common performance management traps:

- We have clear norms and rituals about giving feedback in the moment, for instance debriefing meetings soon after they happen (if not in the meeting itself)
- We are clear about what responsibilities each individual has, what the explicit "brief" is for each responsibility (what outcomes, by when, subject to what constraints) and when responsibilities change. We use a tool called "responsibility maps" to capture an overview of responsibilities for each individual in the firm
- We separate evaluation of outcomes from evaluation of people, in order to be able to separate out a clinical assessment of "project X fell short of our goals" from an assessment of whether, for instance, the manager of the project fell short of what could reasonably be expected of her in the circumstances
- We take time to understand why outcomes happened, investing in more and less formal post-mortems so that we can align what we're learning about people, team collaboration, processes, etc. as various good and bad outcomes happen in our work. We make sure to connect these learnings to our understanding of the strengths and weaknesses of people
- We create space to think at pattern level about how people are developing, with 1:1 development meetings approximately monthly that are intended to focus on learning rather than problem solving about the work at hand
- We align on performance assessment vis-à-vis specific responsibilities before the developing semi-annual reviews, so that the review doesn't have to provide new information about how this project or that area of operational responsibility went and can focus on the higher-level pattern
- We grade performance on an absolute scale, and don't impose any specific distribution of individual performance ratings. We don't manage to a fixed bonus pool, given our understanding that if in fact every individual were performing far above expectations, there ought to be plenty of money to pay the correspondingly high bonuses. We separate out a rating for firm performance that adjusts every individual's bonus up and down from individual ratings – these two factors then combine to determine the actual payout – so that messages are clear, and so that generosity in good years or the need to be austere in bad years doesn't undermine the integrity of performance assessment

In this context, our twice-a-year reviews aim to achieve four objectives:

1. Share an overall synthesis of how the individual is performing in their role. At the end of the year this connects to a clear articulation of the impact of individual performance on their variable compensation: at a baseline level if they are meeting expectations, and anywhere from 50% above to 50% below this baseline based on exceeding or falling short of expectations
2. Offering our best understanding of the patterns in the person's skills and behaviors that have most impacted their performance
3. Offering the best guidance we can about how the individual can best develop going forward
4. Connecting our assessment of the trajectory of the individual's growth to his or her increase in compensation from year to year (usually at the end of the year; occasionally off-cycle at mid-year if there is such a large gap between the work someone is doing and how they are being paid that it feels essential to adjust them upwards before year-end)

The first and fourth objectives are important to make clear and transparent, but they take up about 5% of the "real estate" of the written review. The other 95% is a person-specific, thoughtful narrative that looks backward and looks forward, for which the objective is to be of service to the person being reviewed. (For more discussion of some of these elements of reviews, see a post from last year on Beyond the Ritual of the Annual Review.)

What Makes a Review Useful

To be useful to the individual being reviewed, a narrative review needs to help them see something they otherwise wouldn't see or wouldn't see so clearly – or needs to help them better take action on something that they see already. Because we have so many other mechanisms to help people see what's going well and poorly in the course of doing their work, reviews should aim to help people see patterns more clearly, synthesize better why and how these patterns are important, and figure out how best to be purposeful about their future development.

In many cases, reviews usefully connect the dots in terms of the progress an individual is making and the patterns behind that progress. For instance, take the following passage, edited lightly to be understood by an outside reader who wouldn't know the specific examples called out in the original and to preserve confidentiality:

1. Agenda shaping. Of the areas we called out in your last review as areas of development, you have made the most decisive progress on agenda shaping. This progress is exemplified in the way you navigated your work at [entrepreneurial venture X], tacking away from the focus on management systems during a period when hiring needed to be the central focus, finding a way to tack back towards this agenda when the right conditions were in place to complete that work, and helping weave a focus on leadership development and the founder's role into the core of our work. This represents a step forward from where you were, for instance, at [corporate client Y] last year, where you shaped an agenda within a specific functional area, but didn't connect what you were seeing and experiencing nearly as strongly to setting agenda for our work with the leadership team as a whole.

This kind of dot-connecting is useful in helping someone who is developing well consolidate progress, keep focused on what's beginning to bring them to the next level of development and establish a crystal clear visualization of "what good looks like."

At other times, the relevant dot connecting is in getting underneath a pattern of what an individual should do differently. Another lightly-edited example:

2. Reallocate effort from inputs to outputs and outcomes. You consistently produce intermediate work product of very high quality, across a spectrum from more "raw" products like agendas and meeting notes to more "baked" intermediate products like working drafts of an investor deck. These products are certainly valuable, and you've mastered the ability to create intermediate products in an exceptionally efficient way. Notable as these strengths are, looking across several months of your work, a pattern stands out that the total amount of effort across all of these intermediate products appears large, in ways that have almost certainly crowded out bandwidth that

could have been placed against larger, more difficult questions of how to achieve critical outcomes. One important part of this involves pulling yourself up above the level of deliverables, to the more senior work of finding the leverage points to influence the end outcome with greatest economy of effort. For instance, figuring out how, in the context of [client CEO's] limited bandwidth, we can tangibly impact the way his leadership team comes together and the way this team sets its collective agenda could be the pivotal factor in whether we achieve the impact we aspire to. Finding the right moves to influence these outcomes relating to the senior team – some of which may require relatively little time — could make a greater difference than many, many other things on which you might expend a great deal of energy.

The examples above focus on helping someone see vividly where they are – e.g., the fundamental step forward on agenda shaping; the need to reallocate focus from just being productive in the context of tangible deliverables to being impactful in situations where "what to do" is far less clear. There's value as well in coaching regarding how to move the needle going forward. The following passage provides an illustration:

3. Situation assessment. One important driver of your effectiveness will be the capability to quickly and accurately take stock of each situation you're in: what are the most important goals to advance; where are each of the other individuals involved focused and why; what are the most significant ways you can contribute and what will those contributions require; do you need to influence the direction things are taking; what is most important to be watching for as the situation unfolds? With practice, this "orient" step (to use the vocabulary of Colonel John Boyd's OODA loop: observe, orient, decide, act) can become a distinct mode you go into regularly, stepping above the modes of observing on the one hand or acting on the other and taking a structured view of what's going on, what's at stake and what outcomes matter. Given that your work will generally include a range of competing objectives – many of them changing and evolving as conditions change – the capacity to orient quickly and determine the implications for both actions now and potential actions later will be an important productivity skill.

Of course there is nothing magical about the timing of a semi-annual review as a vehicle for sharing the kinds of insights these three examples above illustrate. Discussions like these could be useful to a member of our team at any time of the year, and in fact we invest in 1:1s between each individual and a senior person who plays a "development lead" role for that person approximately once per month precisely for that reason. The discipline of semi-annual reviews, however, ensures that a clear picture comes together at about the right frequency to ensure that the individual and the firm always have a good, relatively current synthesis of strengths, development needs and career trajectory. People know where they stand, and have more fortitude with which to face the inevitable ups and downs of a day-to-day environment in which the work is *hard* and there's lots of feedback because they have no reason for insecurity about where they stand in terms of the firm's overall view of their capabilities and performance.

Do We Recommend Our Practice to Others?

Sharing this detailed view of what we do and why we do it illustrates a broader point about how

suspect the idea of best practice should be. In a sense, I believe that our approach to reviews is a best practice. I've heard over and over again how moved people have been by the quality and quantity of thinking in reviews about what makes them tick and how they can best develop. We see proof from day to day of how much people refer to and act upon the themes discussed in their reviews, and of the difference this makes to our work in teams and with clients.

For all those positives, we wouldn't recommend our practice to everyone. It takes a lot of time and skill to synthesize useful insights in narrative form. Companies that don't invest in the supporting disciplines that make this approach work for us – such as the quantity and quality of in-the-moment feedback or the care to discuss performance at the responsibility-by-responsibility level before crafting the review – probably wouldn't get the same benefits from the time we invest in our process. Firms that believe they need to "grade on a curve" to hit a fixed bonus pool number will inevitably muddy their messages about individual performance by the need to fit the curve, and should acknowledge what they're doing. It's likely not plausible that Deloitte or Accenture can, at their scale, manage a system like ours – and eminently sensible that they don't try.

Even if some of our specific choices may not translate outside our immediate context, the way of thinking behind those choices does have broad applicability. The principle of separating out assessment of outcomes from assessment of people is even more critical at GE than at Incandescent, given that the causality of outcomes reflects even more complexity and interdependence in a mega-company than in a small firm. The importance of being able to connect a clear picture of "why am I getting the outcomes I've been getting" to a clear picture of "how can I get better" applies to any kind of work, to a musician or an athlete just as much as an Incandescent team member, and equally to a call center team leader or to the Executive Vice President of a megacorporation.

At a time when so much of the debate about performance is about the very real need to throw out the very dirty bathwater many companies have been letting stand for years, it feels important to talk about the "baby": the ability to have a real, human, unvarnished discussion about performance, the patterns driving performance and what path to improvement has the greatest impact and the greatest potential to succeed. These conversations are, like everything in business, subject to the human frailties of incomplete understanding and imperfect judgment. We honor each other and advance both our team and our business when we refuse to shy away from the critical work of establishing the best understanding of individual performance we can.