

Performance Evaluations: How Is This Still a Thing? March 24, 2016 draft

This paper was researched and written by:

Correen Bales, Sandoval Regional Medical Center, Rio Rancho, NM

Clark Carboneau, Presbyterian Health Services, Albuquerque, NM

Hal Moran, Bridgeway Behavioral Health, St. Charles, MO

Mike Stoecklein, Institute For Enterprise Excellence, Lakeville, MN

Our intent is to submit this paper for presentation at the Annual International Deming Research Seminar in 2017.

Table of Contents:

1 Executive Summary

2 Dr. Deming's comments and reasoning on the topic of performance evaluations

3 Key points from other thought leaders and authors

4 Some companies that have discontinued or fundamentally changed their performance evaluation systems

5 Our research and findings

6 Performance Evaluations: How Is This Still a Thing? Our thoughts and conclusions on the current state and trends

1 Executive Summary.

Our paper begins with a summary Dr. Deming's thoughts and comments on the performance evaluation system, including key points from his 1986 book *Out of the Crisis* and 1993 book *The New Economics*.

Several notable thought leaders and authors have also written extensively on the topic of the annual performance evaluation system. We include the key points from several authors including Peter Scholtes, Mary Jenkins, Tom Coens, Alfie Kohn, Samuel Culbert and Brian Joiner.

Recent examples from large and notable companies that have fundamentally changed their annual performance evaluation system are highlighted. These companies include: Accenture, Adobe Systems, General Electric, and Microsoft. It is too early to tell if these changes are permanent, nor is it clear that the changes also address the problems with evaluator bias and variation.

Even though prominent thought leaders and authors have clearly outlined the problems associated with the performance evaluation system, the system still remains in place in most organizations. We surveyed 41 individuals from 25 organizations (primarily healthcare) about the current state of their performance evaluation systems. The survey process included representatives from executive management, management, front line staff and support functions (including lean, process improvement, quality and human resources). Here are some of the questions we asked:

A. Describe your current performance evaluation system. How does it work?

B. If you no longer have a performance evaluation system, please describe how you discontinued it.

C. Have you ever tried to discontinue (or fundamentally change) your performance evaluation system? If so, describe your experience.

- D. What is the purpose of your performance evaluation system?
- E. Is your performance evaluation system designed to affect: 1) aligning strategy, 2) promoting cooperation between departments and people, and 3) improving the performance of work systems. If so, how?
- F. If systems drive behaviors; what behaviors are your performance evaluation system driving?
- G. If you were to discontinue your performance evaluation system, what do you think would happen?
- H. Why do you think your performance evaluation system persists?
- I. If you have been the recipient of or the provider of performance reviews, how did you feel or act afterwards both in the short term and long term?

We conclude with some of our key findings on the question about why performance evaluation systems are still “a thing.”

2 Dr. Deming’s comments and reasoning on the topic of performance evaluations

Dr. W. Edwards Deming’s consulting practice became very busy after he was briefly featured in a 1980 NBC documentary “If Japan Can, Why Can’t We?”¹ He was in high demand, getting requests for what management should be doing, and his 4-day seminars were seeing increased attendance. He provided (somewhat reluctantly) a list of points to management on what they must do (or not do) in order to survive. This became his “14 Principles for Transformation of Western Management”. These evolved over time. There were not always 14, and the wording of the points changed as Dr. Deming learned more about what management must do (or stop doing). These 14 points were described in some detail in his 1986 book *Out of the Crisis*.²

Point 12 always generated discussion and push-back at the 4-day seminars or when people read his books. Here is the wording for point #12 from *Out of the Crisis*:

- “12a. Remove barriers that rob the hourly worker of his right to pride of workmanship. The responsibility of supervisors must be changed from sheer numbers to quality.
- 12b. Remove barriers that rob people in management and engineering of their right of pride in workmanship. This means, inter alia, abolishment of the annual or merit rating and management by objective.”

Dr. Deming elaborated on the problems caused by the annual performance evaluation system as follows: “It nourishes short-term performance, annihilates long-term planning, builds fear, demolishes teamwork, nourishes rivalry and politics. It leaves people bitter, crushed, bruised, battered, desolate, despondent, dejected, feeling inferior, some even depressed, unfit for work for weeks after receipt of rating, unable to comprehend why they are inferior. It is unfair, as it ascribes to the people in a group differences that may be caused totally by the system that they work in. Basically, what is wrong is that the performance appraisal or merit rating focuses on the end product, at the end of the stream, not on leadership to help people. This is a way to avoid the problems of people. A manager becomes, in effect, manager of defects.”²

¹ <https://www.deming.org/content/if-japan-canwhy-cant-we>

² Deming, W. Edwards, *Out of the Crisis*, 1986, MIT-CAES.

In the late 1980s and until he died in 1993, Dr. Deming described the body of knowledge from which the 14 points were derived. He described this “system of profound knowledge” as four interdependent components in his 1993 book *The New Economics*:³

- Understanding Variation
- Appreciation for a System
- Psychology
- Theory of Knowledge.

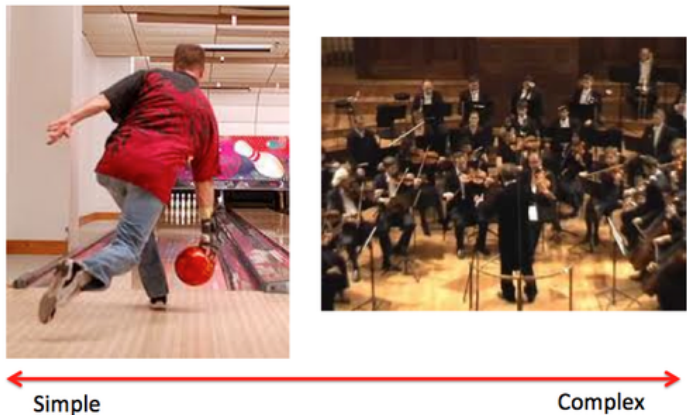
A deeper understanding of these 4 components and their interactions, provide the logic behind why performance appraisals and pay for performance rob people of their pride in work and should be abolished.

Reason 1 - When we understand variation and we have appreciation for a system we will understand that any result (performance, behavior, etc.) comes from 2 factors and their interaction, explained by Dr. Deming as follows;

“You have an equation: $x + (yx) = \text{result}$. Let “x” = contribution of the individual. Let “y” = the system in which the individual works. There is some outcome, 8. It could be \$8,000,000 in revenue or 8 mistakes. We need x. Unfortunately, there are two unknowns and only one equation. Johnny in the sixth grade knows that no can solve this equation for X. Yet people that use the merit system think that they are solving for x. They ignore the other term (yx), which is predominant.”³

So, to focus solely on on the individual (which is what you do with performance evaluations and pay for performance) is not rational. Do the math - it does not work.

Reason 2 - Performance evaluation and pay for performance systems tend to sub optimize the system by driving behaviors that cause people to focus on their individual efforts (results) rather than to consider the larger system of which they are a part. Dr. Deming used a simple example to explain this.



Some systems are simple, like a bowling team, where there is very little interdependence. To know how the system (team) is doing, you merely add up the scores of the individual players.

But many systems are not like bowling teams. They are more complex and resemble a symphony orchestra. The output (great music) is more than the sum of the parts, it is the product of the interactions of the parts. It is not the sum of many solos. The output of a company is not the sum of the parts, it is the product of the interactions of the parts and how they work together. However, western management tends to divide the organization into parts and treat something complex as if it were as simple as a bowling team.

³ Deming, W. Edwards, *The New Economics*, 1993, MIT-CAES.

Traditional performance evaluation and pay for performance systems drive behaviors that force the individuals to manage their boxes on the organization chart, rather than to manage the organization as a system. Systems drive behaviors, we know this to be true. The behaviors that the traditional performance evaluation and pay for performance systems drive are competition between the people, not collaboration.

Reason 3 - When we understand psychology, including what motivates people (intrinsic, extrinsic motivation and the phenomenon of "over justification") we would understand that traditional performance evaluation and pay for performance are some of the destructive forces that squeezes out intrinsic motivation that a person is born with. Dr. Deming used this diagram⁴ to illustrate. Therefore, continuation of a performance evaluation and pay for performance system, or creation of one where it does not exist, contributes to the forces of destruction of the individual.

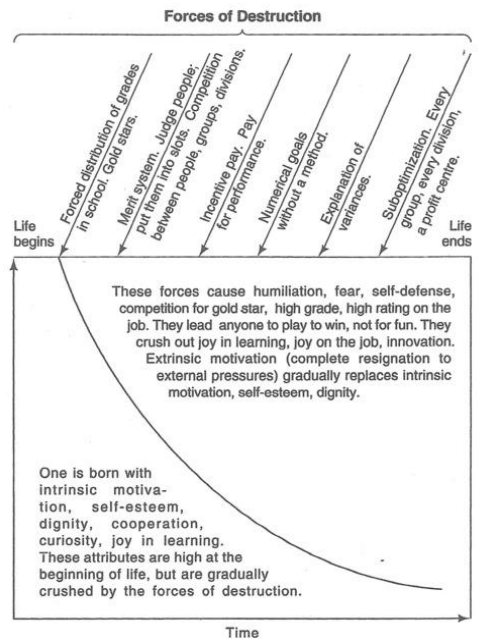


Fig. 10. The forces along the top rob people, and the nation, of innovation and applied science. We must replace these forces with management that will restore the power of the individual.

3 Key points from other thought leaders and authors.

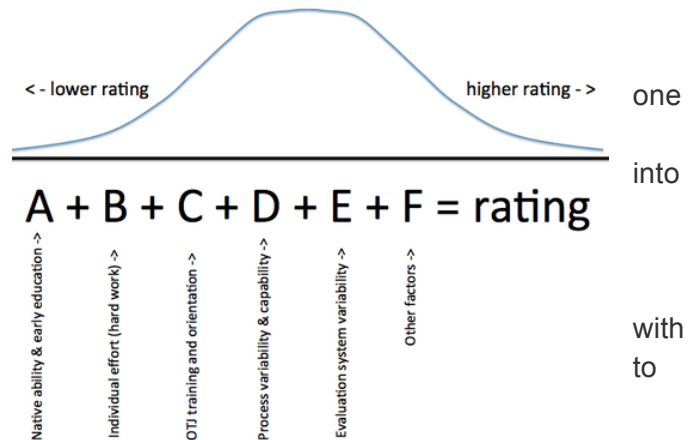
Many other thought leaders have further elaborated on the rationale for eliminating the annual performance evaluation process. Here is a summary of their arguments on the topic.

Peter Scholtes worked closely with Dr. Deming and assisted him in many of his 4 day seminars. Scholtes developed an excellent demonstration about the fallacy of performance evaluations.

First, there is the fact that people get high ratings and low ratings (see figure below). Some are rated below average, some above, poor, outstanding, etc. It's generally a bell-shaped distribution. What are the factors that make up the difference between those that are given high - rating or low rating?

Five factors (listed below - Scholtes allows for a 6th for the workshop participants). In his workshops, he would allow the participants to break up 100 points these categories (not the same #) and assign how much influence they think each factor has on the total rating:

- A – native ability and early education (they just come some good ability and were fortunate to have access good education)
- B – individual effort of this person, how hard they work
- C – training and orientation they received in their job



⁴ From *The New Economics*, Chapter 6, Figure 10.

- D – variability and capability of the processes in which the person works
- E – system of evaluation itself (accuracy or inaccuracy)
- F – other (cited in the workshop: leadership skills of manager, previous experience, compatibility with their manager)

Scholtes then asks the participants who gave each of the factors the highest or 2nd highest points. Here are some of the typical rough percentages received from the audience.

- A – about 25%
- B – about 65%
- C – about 15%
- D – about 40%
- E – about 40%
- F – < 10%

4 of the 5 items are not in the control of the individual. It seems that there's only one factor that is in the control of the individual. The only legitimate basis to rate this person is on "Factor B." Scholtes offered comments on some of the other factors: Factor A – if they are dumb, why did the company select them? Factor C – why is the person rated on the training the company gave? What if the training was poor, inadequate or wrong.

Even if "working hard" (Factor B) is a strong factor, there is some evidence about the effect of the system on this factor. Some systems cause people to work harder, others cause them to not really put in that effort. So, even with "Factor B", the system has an effect. No person can separate out which one is THE factor $A+B+C+D+E + F = \text{rating}$. Can anyone solve this for "B"? (when there are other unknowns?) If you can, you have the right to do performance evaluation. In a nutshell this is Scholtes' case against performance evaluation, similar to Deming describing it as $X = \text{individual}$, $Y = \text{System}$ as described in the previous section of this paper.⁵

Scholtes outlined other arguments against performance evaluation:

Reason 1 Performance evaluations disregard the fact that people work in teams or groups, and this contribution can't separated out. The evaluator could be giving people credit or blame for what came from other people.

Reason 2 Performance depends on the system (the step in the process before you). Who can discern the value added by one step of the process and the preceding steps? No-one. Example, if you write a report, you depend on information from others. The ability to perform is dependent on systems.

Reason 3 Performance appraisal disregards that employees work within variability and instability. The assumption is that the output is all the same. The output is not all the same. Without some understanding of variation, we think there are patterns, but many times we are incorrect. The patterns turn out to be random variation. Managers who don't understand variation won't understand this. Another example would be OSHA reportable accidents, which is mostly due to random variation. How do you evaluate someone in

⁵ Scholtes, Peter, *The Leader's Handbook*, 1998.

a system that is always changing? Managers can make a lot of mistakes (if they don't see that there's a system or don't understand variation)

Reason 4 We assume that there is no measurement error, and therefore assume that the evaluation system is fair, equal and consistent. In reality there is wide variation, from one evaluator to another and between the same evaluator over time. The thing we are least objective about is our own objectivity. Some of the bias is cultural (local, regional, eg physical fitness, politically correct). Systems pressure on the evaluator (e.g. the "book" is made on people, the common characterization of a person, their reputation). It is difficult to see if they are a good worker or not, when the evaluator is influenced by the common view. It is difficult to depart from the common view and focus only on the person. There is also evaluator bias and different evaluator philosophies:

a) One philosophy that some raters have: "all my people are above average, my assumption is that is true unless they prove otherwise."

b) Another philosophy: "all my people are average, if anyone gets above average, it's something they need to earn".

Both are legitimate, which one would you rather be rated on? Answer - "evaluator a" - the easy marker.

Within a company there are easy and hard markers. It's a lottery. It would make more sense if you flipped a coin.

Let's say you can avoid all of these factors (which we know you can't), then what? Is there anything wrong? Yes, Scholtes described other "toxic wastes" of performance evaluation:

Mediocrity – The tendency to pick easy goals, sure winners. If your salary depends on achieving the goals. On the key things, you are going to play it safe, so as to look good. Most people will do this because it's the way the game is played. The organization gets sub optimized.

Not Understanding Capability Management might choose goals that are a stretch, beyond what the system is capable. Everyone trying to do more than what the system is capable of, unless it is studied and improved, it will stay the same. Or only at the expense of some other part of the system.

Performance evaluation creates losers and cynics For any group, half will be above average, half below average or 2/3rd will be below the top third. How does this match people's self image? Most people see themselves in the top 3rd. They have a high self image of themselves. The most objective system will tell half of the group that we are below average. Scholtes asks this question, "do you know the difference between chronically depressed people and the rest of us? Chronically depressed people have a more accurate self image". You'll create losers, or cynics (they will know the system is screwed up).

Other thought leaders and authors who have written extensively on the problems with the annual performance evaluation include: Alfie Kohn⁶, Mary Jenkins and Tom Coens⁷, Samuel Culbert⁸, and Brian Joiner.⁹ Here are some of the key points that they make.

Alfie Kohn "Performance evaluation is an exercise in futility. Evaluation systems are far less accurate than we would like to believe. It is nearly impossible to quantify the performance of many kinds of work. Most

⁶ Kohn, Alfie, *Punished By Rewards*, 1999.

⁷ Jenkins, Mary and Coens, Tom, *Abolishing Performance Appraisals*, 2000.

⁸ Culbert, Samuel, *Get Rid of the Performance Review!*, 2010.

⁹ Joiner, Brian, *Fourth Generation Management*, 1994.

rating systems are only accurate at the extremes (exceptionally good or bad performance). A performance evaluation system may seem precise, but it actually masks many distinctions in the way someone approaches his job. Even when we try to allow for a more qualitative appraisal, that assessment tells us as much about the appraiser as it does the appraisee. Any individual's performance is, to a considerable extent, a function of how he is managed. Assessment of individuals also overlook the extent to which any one's performance grows out of an exchange of ideas and resources with colleagues and otherwise reflects the indirect contributions of the larger system. Even if performance appraisals were adequate to gauge how well people are doing, their effects are usually so destructive that they shouldn't be used anyway. Not only is the fact of interdependence in the workplace ignored, but people are discouraged from cooperating in the future."⁶

Mary Jenkins and Tom Coens "Performance evaluation impedes the reception of feedback, and there is no solid evidence that it motivates people or leads to meaningful improvement. Due to its inherent design flaws, appraisal produces distorted and unreliable data about the contribution of employees. Consequently, the resulting documentation is not useful for staffing decisions and often does not hold up in court. 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. The research decidedly demonstrates that appraisals are ineffective and cause a spate of undesirable, unintended effects. As we excavated a mountain of research on appraisal, we became increasingly amazed that business scholars and leaders have not collectively called a halt to the practice. Unhealthy underlying assumptions about people and work are deeply ingrained in our thinking. If we don't see them clearly and pull out their roots, they will surface again like weeds in the new system design. By digging up and removing the roots of these weeds (making the assumptions of the former system explicit) in the design of alternative systems, we create integrity between what we say and what we believe. Some of the prevailing assumptions that must be surfaced and challenged, and some alternative proposed assumptions are:

- One appraisal process can effectively serve several functions at the same time. Alternative assumption: The multiple purposes of appraisal can be better achieved by separate processes.
- People want to know and need to know where they stand, and appraisals tell people where they stand. Alternative assumption: People want to clearly understand and access the knowledge and information that influences decisions about their pay, promotion, status, and future.
- The organization and the supervisor are responsible for individual employees' morale, performance, and development. Alternative assumption: As healthy adults, people need to be responsible for their own morale, performance, and development, with and without support from the organization."⁷

Samuel Culbert "This corporate sham is one of the most insidious, most damaging, and yet most ubiquitous of corporate activities. Everybody does it, and almost everyone who's evaluated hates it. It's a pretentious, bogus practice that produces absolutely nothing that any thinking executive should call a corporate plus. And yet few people do anything to kill it. Why is that?

- The performance review is all executives have ever known, and they're blind to the damage caused by it.
- Because few managers are aware of their addiction to the fear that reviews create amongst staff, and too many lack the confidence that they can lead without that fear.
- Because HR professionals exploit the performance review to provide them a power base they don't deserve.

- because few people know an alternative for getting the control, accountability, and employee development that reviews supposedly produce— but never do.

Reviewing performance is good; it should happen every day. But employees need evaluations they can believe, not the fraudulent ones they receive. They need evaluations that are dictated by need, not a date on the calendar. They need evaluations that make them strive to improve, not pretend they are perfect. In fact, if firms did nothing else but just kill off this process they'd immediately be better off. When it comes to performance reviews, there's no question that nothing is better than something. That's how bad they are."⁸

Brian Joiner "Organizations tend to load a lot into their performance appraisal system. Some typical uses include:

- giving feedback to employees
- giving direction to employees
- identifying training needs
- fostering communication between manager and employee
- providing evidence for promotion decisions
- providing a basis for compensation decisions
- serving as a defense in legal cases associated with promotions or terminations

How well do performance appraisals meet these important needs? The list alone provides a clue: no single system is ever likely to do a first-rate job on all these disparate needs. When we look carefully, it's evident that performance appraisals conflict with several 4th Generation lessons:

- The implicit assumption is that an individual's performance can be judged in isolation of the system in which he or she functions, something we know misses the point. The evidence is clear. Most problems lie in the system - relatively few due to the individual.
- They fail to recognize variation in perceptions of performance. Have you ever gotten a new manager and overnight your rating dropped from "great" to "not so hot"? Did your performance change that much or was it mostly a difference in perceptions?
- Most deadly of all, performance appraisal systems fail to recognize the extreme importance of cooperation. Many performance evaluation systems actually structure competition among employees by creating a win-lose environment."⁹

4 Some companies that have discontinued or fundamentally changed their performance evaluation systems.

Adobe Systems "In 2012, Adobe replaced its evaluations, after its senior vice-president of people and places, Donna Morris, concluded that employees hated them and they weren't very useful. Under the old system, managers focused on the most recent developments instead of looking at the entire year; they fixated on the past instead of considering what employees should do going forward; and, most important, they came across as antagonistic. Instead, Morris implemented a more informal "check-in" process that takes place throughout the year, with employees receiving feedback on what they're working on at any given moment."¹⁰

Accenture "As of September 2015, one of the largest companies in the world (330,000 employees) did of its employees and managers an enormous favor: It got rid of the annual performance review. Accenture will

¹⁰ <http://www.newyorker.com/business/currency/the-push-against-performance-reviews>

disband rankings and the once-a-year evaluation process starting in fiscal year 2016. Accenture implemented a more fluid system, in which employees receive timely feedback from their managers on an ongoing basis following assignments.

“Accenture is joining a small but prominent list of major corporations that have had enough with the forced rankings, the time-consuming paperwork and the frustration engendered among managers and employees alike. Six percent of Fortune 500 companies have gotten rid of rankings, according to management research firm CEB. These companies say their own research, as well as outside studies, ultimately convinced them that all the time, money and effort spent didn't ultimately accomplish their main goal — to drive better performance among employees.”¹¹

Deloitte “In March 2015, the consulting and accounting giant Deloitte announced that it was piloting a new program in which, like at Accenture, rankings would disappear and the evaluation process would unfold incrementally throughout the year. Deloitte is also experimenting with using only four simple questions in its reviews, two of which simply require yes or no answers.”¹¹

General Electric “For decades, General Electric practiced (and proselytized) a rigid system, championed by then-CEO Jack Welch, of ranking employees. Formally known as the “vitality curve” but frequently called “rank and yank,” the system hinged on the annual performance review, and boiled the employee's' performance down to a number on which they were judged and ranked against peers. A bottom percentage (10% in GE's case) of underperformers were then fired. The company got rid of formal, forced ranking around 10 years ago. But now, GE's in the middle of a far bigger shift. It's abandoning formal annual reviews and its legacy performance management system for its 300,000-strong workforce over the next couple of years, instead opting for a less regimented system of more frequent feedback via an app. For some employees, in smaller experimental groups, there won't be any numerical rankings whatsoever.

“The world isn't really on an annual cycle anymore for anything. Some of it to be really honest is millennial based. It's the way millennials are used to working and getting feedback, which is more frequent, faster, mobile-enabled, so there were multiple drivers that said it's time to make this big change. There's a growing realization that the annual review just isn't a particularly good way to manage people or to boost performance. It leads to a tendency for HR to focus excessively on process over outcomes.

The move by more and larger companies away from annual reviews and ratings is well past due, say management theorists. Years of research, from both business school professors and neuroscientists, has found that the practice is ineffective at boosting performance, actively alienates employees, is based on a flawed understanding of human motivation, and is often arbitrary and biased. People simply don't fit neatly (pdf) on a bell curve. It ends up being an exercise in paperwork and bureaucracy instead of an agent of change.”¹²

¹¹ <https://www.washingtonpost.com/news/on-leadership/wp/2015/07/21/in-big-move-accenture-will-get-rid-of-annual-performance-reviews-and-rankings/>

¹² <http://qz.com/428813/ge-performance-review-strategy-shift/>

Microsoft “Microsoft did away with its rankings nearly two years ago, attracting particular attention since it had long evangelized about the merits of its system that judged employees against each other.”¹⁰

However

“The new systems at companies like Accenture, Deloitte, and Adobe certainly seem friendlier and more streamlined than the past approaches. Lately, though, companies and researchers have drawn attention to another problem with performance reviews—one that these companies’ new practices may not be addressing and could, in fact, be exacerbating. Data shows that all kinds of personal quirks and biases, both conscious and not, influence our appraisals of other people. (If you’re not sure that you believe in unconscious bias or think you’re immune, try these tests of your own racial and gender prejudices.) These prejudices, which are sometimes collectively referred to as “the idiosyncratic rater effect,” influence how managers think about—and describe—their employees. Studies suggest that more than half of a given performance rating has to do with the traits of the person conducting the evaluation, not of the person being rated.

“Anyone who has been given a negative performance review by a mean-spirited or incompetent boss will be familiar with this effect, but its implications go beyond individual dispiritedness. As James Surowiecki wrote last year, in a piece about the gender imbalance in Silicon Valley, “In a recent study of almost two hundred and fifty performance reviews, the tech entrepreneur Kieran Snyder found that three-quarters of the women were criticized for their personalities—with words like ‘abrasive’—while only two of the men were.” After Ellen Pao, a former junior partner at the venture-capital firm Kleiner Perkins Caufield & Byers, sued the firm for discrimination, in 2012, it emerged that her performance reviews at the firm had described her as having “sharp elbows,” and being “territorial” and “not a team player.” (Pao lost the case, but the evaluations reflected poorly on Kleiner Perkins.)

“Organizations, need to make sure their performance-evaluation process attempts to minimize bias—which often involves collecting more, not less, information and documentation, and having more, not fewer, people weigh in on an employee—in other words, ‘the opposite of what organizations like Accenture appear to be doing,’ The frustration with performance appraisal is certainly understandable, but if alternatives become more subjective and less attention is given to reducing bias, there is potential for serious problems down the road.”¹⁰

5 Our research and findings: our methodology and our findings

Our study included conversations with 41 individuals from 25 organizations. The roles of the participants were: Executives 14, middle management 11, front line 3, PI/Lean 10, Human Resources 2, and other (quality) 1.

23 of organizations represented in the study have a formal performance evaluation, 2 have discontinued their system. Of the 23, most (19) have an annual process (although within this group at least 6 have some managers having check-ins more often – weekly - quarterly). Three organizations apply their process every 6 months, one has a quarterly cycle.

Of the two that have discontinued formal performance appraisal system, both recognized the system as a wasteful and unproductive effort that drove strong negative behaviors. One of these organizations has

developed a system of frequent check-ins from the CEO on down to the front-line level, connected to strategy deployment and leader, manager, supervisor and staff standard work systems. This is a healthcare organization that discontinued using the Joint Commission on Accreditation For Healthcare Organizations (JCAHO) several years ago. They have used the Baldrige framework as a guide, and have passed reviews at by the the State review program and have therefore eliminated the factor of “JCAHO Accreditation” as a reason to retain the annual performance evaluation system.

The other organization that has eliminated the annual performance evaluation system is is a multi-specialty clinic. They have not yet fully designed a system to provide the regular check-in and feedback process at all levels of the organization that they would like to use as a substitute for the annual performance evaluation system.

12 of 39 individuals from orgs with performance evaluation system mentioned the inclusion of a “self-assessment” by employee (4 of these individuals mentioned that their boss depends on them to do this self-assessment – if they didn’t do it, they would not get an evaluation). These 12 people represented 9 of the 23 organizations. This does not mean that the other 14 organizations do not include the self-assessment part of the process, but rather this is the number of individuals who mentioned the self-assessment process as a part of the system.

21 of the individuals from the 23 organizations with performance evaluation systems described some type of rating process. One person identified a 2-level scale (either met, or not met expectations). Three individuals identified a 3-level scale (exceeds, met or not met), Two identified a 4-point scale (exceeds expectations, meets, developing, not meets), Four identified a 5-point scale, One identified a 10 point scale, and one described a “9 box” rating (potential - low to high - on one axis, and performance –low to high – on other axis). Employees are ranked and rated on two scales (potential and performance) and those in the lower left-hand box are designated as “Up or Out” (either they move up, or they are moved out of the organization). Those in the upper right-hand box are both best performers and best potential (advancement material).

Individuals from two organizations mentioned a process by which all of the evaluations in the organization are “calibrated” – to force the data into a distribution, primarily for the purposes of administering pay for performance. The employees in these organizations are not aware of the results of the evaluation until after the calibration or committee process. Three individuals described their process as series of three evaluations: self evaluation, evaluation by supervisor, final evaluation at the executive level by an individual or committee.

The 39 interviewees from organizations that have performance evaluation systems **mentioned** components of their performance evaluation systems summarized in the table below. (Note: the survey did not prompt interviewees for specific items, rather we observed how often they identified these components as a part of their performance evaluation system without prompting for specific items):

Component	# Individuals	# Organizations	Percent (of 23)
Focus is on organization goals	15	12	52%
Focus is on the job description	12	10	43%
Mentioned wide variation in application of the system	12	10	43%
Focus on company values	10	9	39%
Includes action plan for improvement	10	8	35%
Tie pay to performance - yes	9	7	30%
Tie pay to performance - no	7	5	22%
Focus on results	6	6	26%
Focus on competencies	6	6	26%
Focus on behaviors	6	6	26%
Focus on development	4	4	17%
Focus on advancement	3	3	13%
Focus on feedback	1	1	4%

Question #3 – Have you tried to get rid of it or fundamentally change it?

No (11)

Managed by OD department, no interest to change 1

Can't because of JCAHO 1

Not a high priority, it's "not THAT bad" 1

Controlled by corporate, can't change 1

Yes (15)

Too arduous, time-consuming, tried to simplify 5

Away from "pay for performance", the focus was all about pay 5

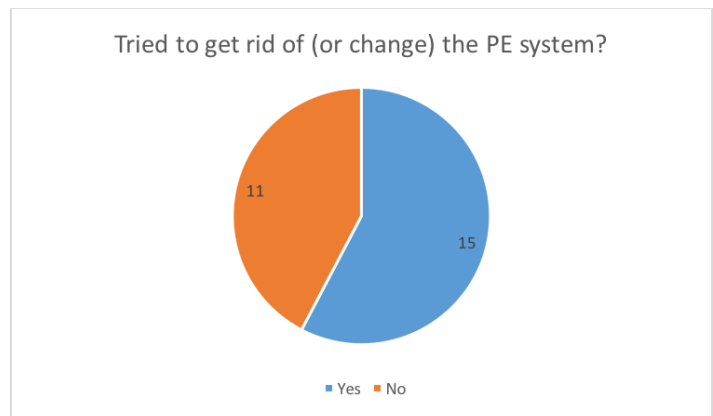
No longer weight or numerically rank 3

Added pay for performance or incentive 3

More frequent cadence 2

No longer forced distribution 1

Not assess competence 1

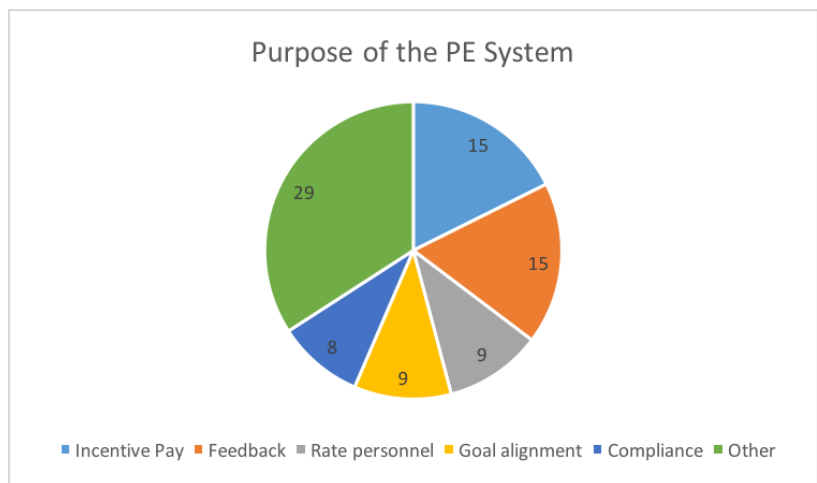


Adding improvement/lean work 1
 Tried “pass/fail” 1
 More focus on alignment discussion 1
 More flexibility for managers 1
 More focus on feedback 1

Have plans to change (7)
 More focus on feedback or coaching 3
 More frequent cadence 3
 Focus on values 2
 More “position” focused 2
 Include PI/lean work 1
 More meaningful and pertinent 1
 Simplify 1
 Less “check box” 1
 Focus on ideal behaviors 1

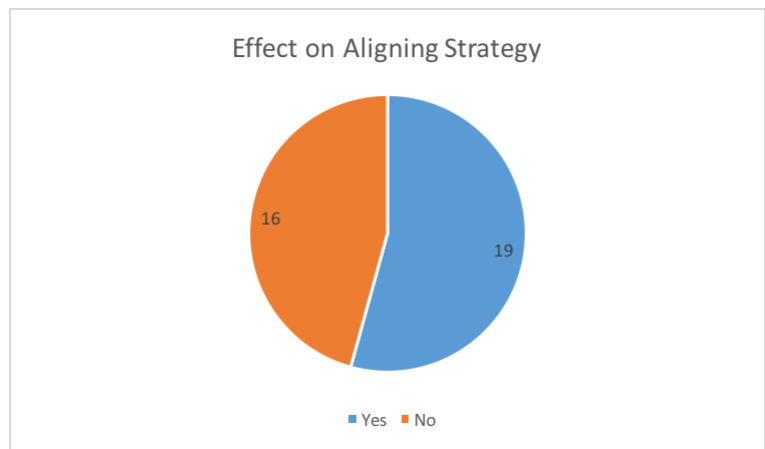
Question #4 – What is the purpose of your performance evaluation/ system?

Pay raises/incentive 15
 Feedback 15
 Rate personnel/high performers/low performers 9
 Goal setting/goal attainment/alignment 9
 Check-box/compliance with rules/regulatory/document for personnel file 8
 Development/growth 7
 Conversation between supervisor & employee 4
 Communicate expectations 3
 Coaching 3
 Record accomplishments 3
 To affect behavior 2
 Plan for low performers 2
 Motivate employee/drive engagement 2
 Recognize employee 1
 Assess competencies 1
 Mobility options for high performers 1



Question #5a – Performance evaluation system’s effect on aligning strategy

Yes 19



Comments:

- Gives people a clear understanding of what their goals are.
- Connected to true north.
- Goal setting, goal cascading
- Every team member, salary and hourly, goes through the same performance evaluation. This way it reinforces our culture and makes sure all team members are aligned with our values.
- Doing better, organizational goals trying to hit the alignment of strategic things (eg engagement in lean management, yellow belt)
- This is a point of consideration for the new system that is being developed. The desire is to reinforce the strategy of the hospital and the importance of the role each employee has in attaining this.
- Aligns accountability
- Management evaluations link to organizational goals & strategy – Employee Evaluations should include goals linked in the same way. Not sure we are consistently doing this.
- At the high level, a conversation around alignment of direction (supervisors and above).
- The process requires that each employee sets 3-5 goals that align to strategy- generally, the system is structured to require a goal on customer/pt experience, clinical quality/safety outcome. Goals are cascaded from the CEO through the system so there is alignment
- The format of the evaluation is setup to align with strategy. By making everyone's goals line up to the strategic initiative.
- It enables the employees individual goals to align with the strategic goals of the organization.
- Part of the performance review is the employee's goals which should be aligned to senior leader goals which are aligned to strategy
- It helps employee set personal goals as well as contribute to the unit achieving their strategic goals.
- It makes sure the employees are working towards the same goals which roll up to the strategies defined by senior management.
- It aligns to strategy by creating guidelines and protocols. It does this by aligning the middle management with the upper management initiatives and then middle management develops guidelines and protocols for the front line staff
- I think the system does attempt to align individual goals to strategic goals. It does this by having the senior management set the organization direction and then cascading down through middle management to the front line staff. Metrics are established on the board scorecard and monitored monthly throughout the year. However for the most part the employees are only evaluated once at the end of the year.
- It has an effect on aligning strategy because the employee's goals are set to support the manager's goals which are set to align with the organizational strategies in the levels above.
- Link to organization core values, goals. Clear performance expectations. What needs to be done to be success. Helps department to deliver on commitments. Regular check-ins to correct work early. Summary of conversations during the year.

No 16

Comments:

- focus on how well the person does that job
- Gives input to align strategy. Assess emotional intelligence (EQ), looking for people with strengths for key strategies.

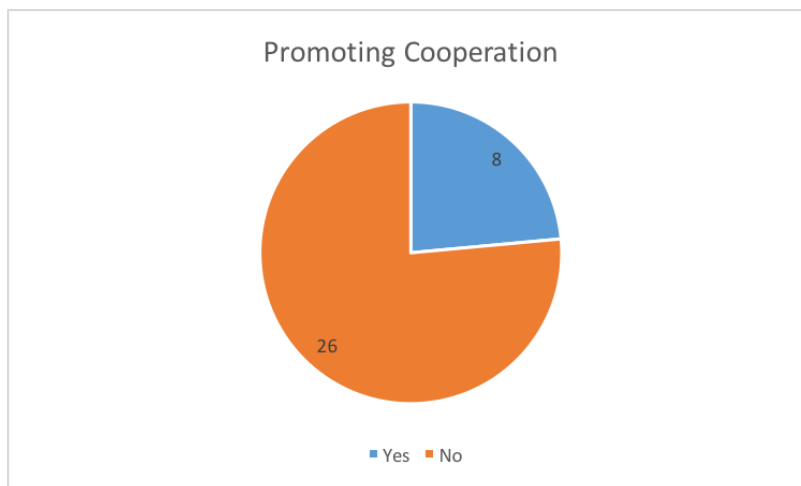
- Minimal to present. The goal this year to change that and have frontline staff help align the organization with the true north goals.
- To employ skilled professionals.
- In the catch ball process, the goals are there, which feeds into the performance evaluation. it is going in this direction at our organization, but not there yet. How does this work at the front-line level, it is not going on yet.
- It depends on how a manager implements it, creates line of site. Manager dependent. The system just loads the job profile. It is too generic. Don't have systematic way to cascade the goals.
- Intended to align to strategy - strategic plan (e.g. improve financial viability, and improve access) can incorporate this into the conversation, but it is not tied directly.
- It's a tool, used well by boss. It is dependent on the manager. took a while (2-3 years) to see that we were being strategic. Outside of the department, there is high variability. Only a handful are really good at using the tool.

Question #5b – promoting cooperation between people & departments

Yes 8

Comments:

- There is a behavior under “people” category (upstream, downstream), it is not a strong connection
- multi-rater process.
- This is done via our values (People, Leadership). Any Silo building is discussed in the PE.
- Internal departmental communication is part of the evaluation



- Employee evaluations score on the Value of Teamwork. Executive level does not. Adding Key behaviors will help to reinforce cooperation between departments.
- Help staff see the bigger picture. Builds camaraderie.
- Because the goals are aligned to high quality care and patient experience.
- Competencies of collaboration, customer focus

No 26

Comments:

- Not directly. Old system did the opposite of cooperation, incented to score people high (could have been someone you wanted to get rid of).
- Leadership - to the extent they have to collaborate (but it's not an intentional design) most managers - focus is “personal goals”, as you go up it is more organizational goals
- At leadership level, it is possible that people “band together”, but this really doesn't happen.
- A couple of questions address this but very little substance.
- Based on the “how” you do things, could achieve a result, but if you didn't involve anyone else. Doing it “with” other people is built in. Organization is changing their values. Some people do 360s, it is optional.

- None really, except more incentive to come together and work together, conversations (at the top of the organization).
- Not sure, If it does, it is not engineered at a system level
- I would say no because everyone sets their goals independently. I have never seen cross functional goals set.
- The calibration portion of the evaluation increases conflict between departments and it takes away from evaluating the employee and then compares against someone with different expectations.
- It increases conflict between departments and within departments.
- I think it is meant to promote cooperation; however the calibration process which pits employees from one department against other departments does not promote cooperation. By design the calibration process fosters competition between and within departments. Each department manager wants their employees to be rated the highest so sometimes the manager has to discount the other department's employees.
- I think the senior managers believe the system promotes cooperation by aligning individual goals to strategy. However, I think in some cases the system fosters competition between staff instead of promoting cooperation. This occurs in at least a couple of ways. The first is the Calibration process which ranks employees against each other in different departments. The second is the system only allows so many high ranking employees, although the specifics of how many is not clear to me. This means department managers downplay the accomplishments of other department staff and in some cases word gets back to the employee of the discussion. This causes staff unrest and bad feelings between departments. Why would staff want to assist other departments and then get put down.
- It has the opposite effect of promoting cooperation because the main use is to stratify pay for performance.
- - You have different departments vying for those dollars. It has resulted in competitiveness rather than cooperation.
- No because the evaluations are done in silos
- No connection. It's a personal information reporting system
- Only time it would have impact is when there are specific concerns or ideas brought up in the conversation that would impact collaboration (e.g. concerns with interactions with other departments or other individuals). Can be used to advise and direct the employee on what they can do.

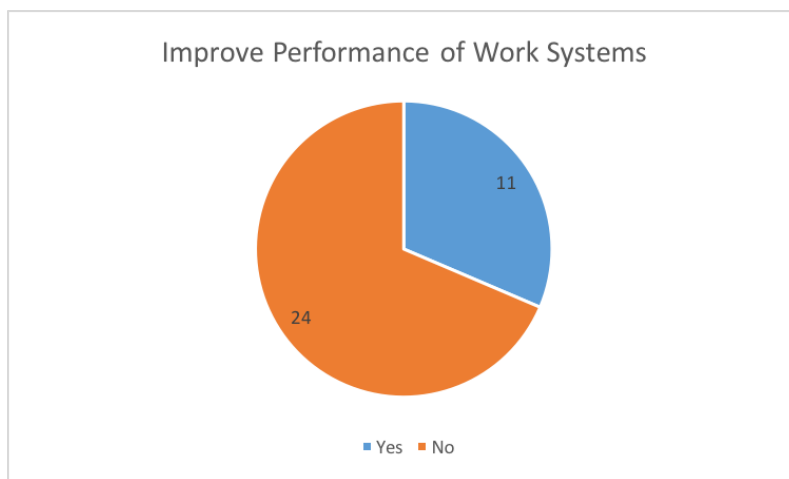
Question #5c – improving performance of work systems/processes

Yes 11

Comments:

- Promotes regular review of the processes and expectations of staff.
- The conversation about the work processes, not the first time people are hearing about issues. Addresses behaviors, not so much actual processes. A manager who has done this a lot might be able to pull in this aspect, but not at the beginning.

- If it gives them opportunities for mentoring and developing.



- We ask - what about our systems needs to be improved or what benchmarks need to be adjusted. identify the gap in individual performance, agree and work on a plan, the customers can benefit from this. “managing people up” someone in organization recognizing someone that did something great. Acknowledging, recognizing people.
 - It promotes improving processes in service to the key results based in the goals.
 - There is a specific question on how the individual’s work impacted customer experience.
 - It helps drive engagement from the front line that are the process experts.
 - Based on the results of goals and their opportunities for improvements the staff makes changes which ultimately improve our service to our patients
- This has the closest tie. Where we are on our lean journey. Managing daily improvement, suggested solutions, committees working on A3s.

No 24

Comments:

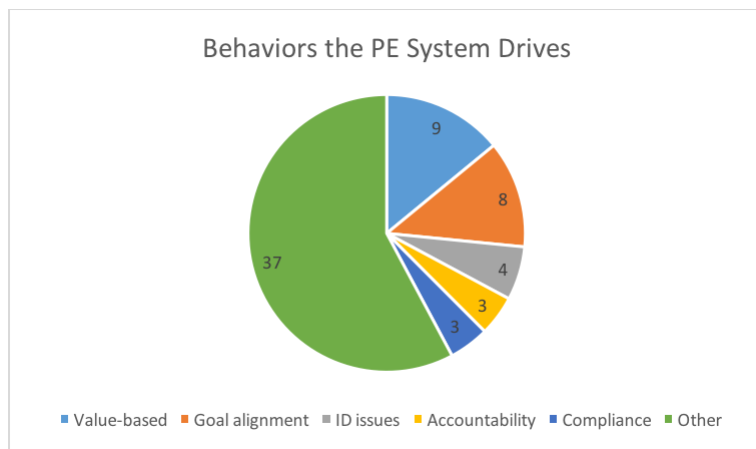
- The PE itself does not improve processes. However, it does start the conversation of any barriers in the way of success for the team member.
- Leadership - depends on how the goals are tied to true north & patient value. It’s not strong
- That’s a goal. I think that the evaluations are completed a little different for the departments. Perhaps could be done better as a system rather than specific departments.
- “Can’t connect the dots between improving work systems and the appraisal process.”
- Performance review doesn’t create the change, the conversation makes the change.
- Only focused on staff.
- The system is done on the individual level not aligned to another group or process.
- If there are specific strategic goals focused on customers and then the employee goals were aligned to the strategic goals then it could have an effect but it would difficult to prove the impact.
- It diminishes it. Because it reduces engagement because people are unhappy after their reviews.
- Employees discuss pay rates even though they were instructed not to. So the system limits the number of good workers by design. In other words there can only be so many often exceeds within a department even though more folks are deserving. The system inherently creates minimalist work environment, i.e., there is no incentive to work harder.
- Staff would contribute more if the system were changed to clearly award system improvement. A number of years ago management was incentivized to attend green belt training. All management attended but most did not complete an improvement project because only attending the class was required. If the system laid out incentives for improvement the organization would have more processes improved.
- I don’t think it is targeted towards processes it is targeted towards individuals.
- In general no unless the goals are specifically about improving systems & processes
- No visibility for this. Hiring manager doesn’t have access to this.
- Doesn’t really affect this.
- Depends on the manager that uses it.

Question #6 - If systems drive behaviors, what are the behaviors that your performance evaluation system are driving?

Value-based behaviors – 9

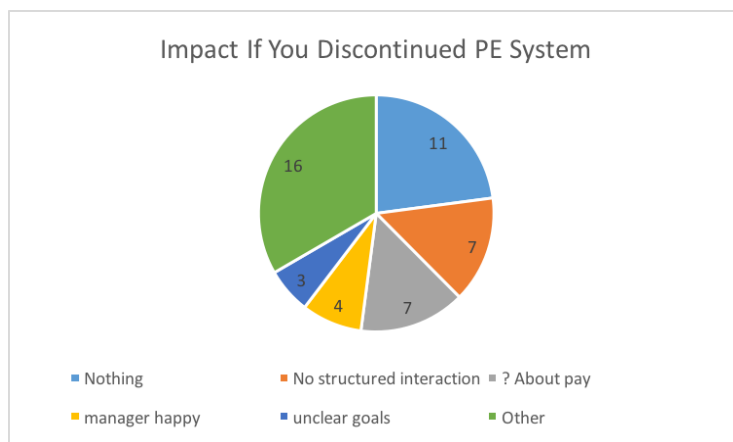
Focus on goals/Alignment/expectations/results – 8

- Look for waste, inefficiencies, issues - 4
- Accountability – 3
- Checking boxes/compliance/paperwork/waste – 3
- Competencies - 3
- Complacency/Doing what they did yesterday – 3
- Drives customer satisfaction/focus - 3
- Get a pay raise - 3
- Silos/focus on the parts, not the whole – 3
- Gaming the system/Low-balling the goals/self-promotion - 2
- Ranking & Rating People – 2
- Annoyance - 1
- Attendance - 1
- Camaraderie - 1
- Career development - 1
- Competition - 1
- Community relations - 1
- Development/growth -1
- Drives sales - 1
- High performers “throttle down” - 1
- Internal departmental communication – 1
- Instill sense of pride - 1
- Personal connection between leader & staff - 1
- Recognition - 1
- Resentment – 1
- Structure - 1
- Track accomplishments – 1
- Work performance - 1
- Write up poor performers – 1



Question #7 - If you discontinued your current performance evaluation system, what do you think would happen?

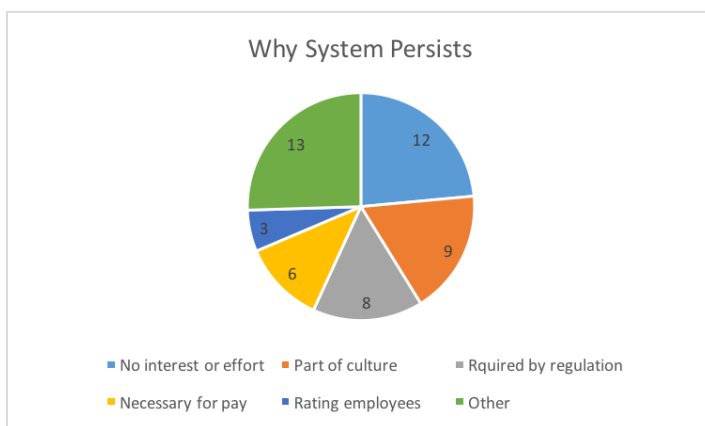
- Nothing would change – 11
- No structured interaction between managers/staff, no personal feedback – 7
- Questions about how to give pay increase/reward top performers – 7
- Manager would be delighted/relieved (less non-value-added work) - 4
- Clear goals would not be set & understood - 3



- those who develop staff would continue to meet periodically to give input – 3
- Low performers would get moved around the organization - 2
- Motivation would decrease – 2
- Save lots of time – 2
- Over time would be erosion in performance – 1
- Job satisfaction would diminish – 1
- Increased collaboration & teamwork – 1
- Customer service would improve – 1
- Accountability would diminish – 1
- Chaos – 1
- Make a feedback system – 1

Question #8 - Why do you think your current performance evaluation system persists?

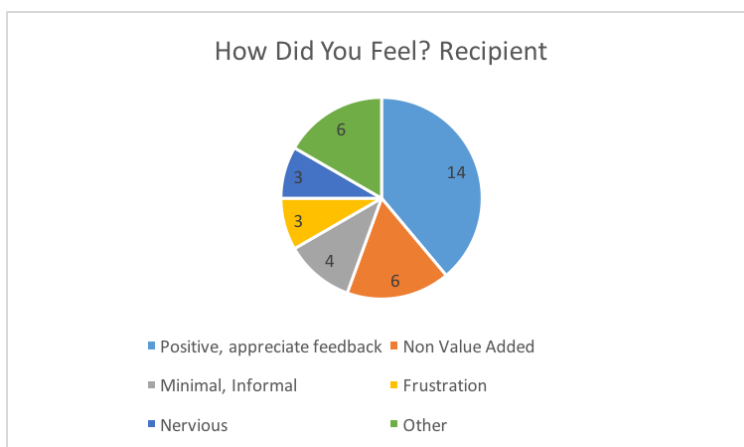
- There is no effort or interest to change it or get rid of it/don't see it as a problem – 12
- it's been around a long time/part of the culture – 9
- It is required (or perceived requirement) by Joint Commission, government – 8
- Think it is necessary to tie pay to performance - 6
- It is needed to rate level of performance for staff in a standardized format – 3
- It saves money (reduces the amount of employee raises) – 3
- It recognizes people for their accomplishments/ID top performers – 2
- It helps people to see where they need to improve – 2
- it is needed to create incentive among staff – 1
- It gives visibility to where future leaders will come from – 1
- Because you have to have one – 1
- It is needed to have a regular conversation between manager and staff – 1
- It is supported by senior management – 1
- It does a lot of things (goals, feedback, development) – 1



Question #9 If you have been the recipient of or the provider of performance reviews, how did you feel or act afterwards both in the short term and long term?

Recipient of:

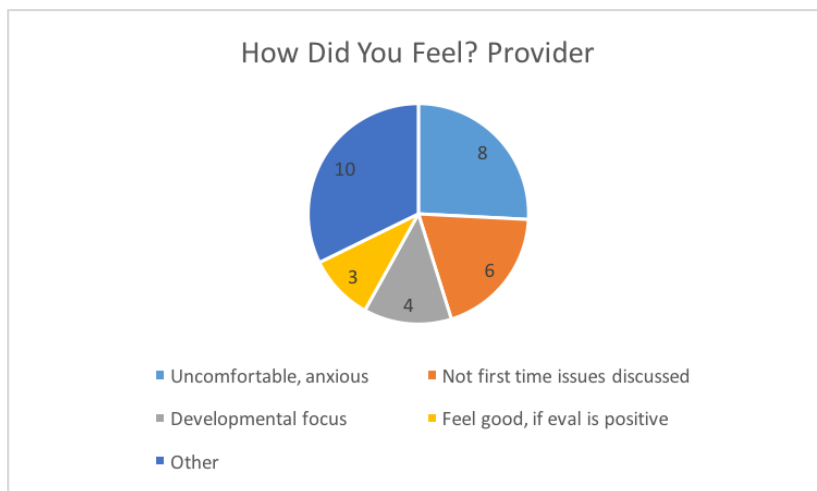
- positive, appreciate the feedback, helps me look in the mirror, helps with my career, motivating - 14



- felt like it was not value added (if reviewer is not really engaged, not prepared or disinterested) - 6
- Informal process, talk all the time, minimal impact on feelings - 4
- frustration - 3
- nervous, skeptical - 3
- awkward (don't like to be recognized or "set apart") - 2
- initial "feel good" experience is short-lived when commitments & promises are not kept - 1
- felt unprepared, need to work on documents to better "self-promote" myself next year - 1
- relief (would prefer "on-going" not annual) - 1
- ambivalence - 1

Provider of:

- Feel uncomfortable, anxious, postpone and put it off if I anticipate there will be issues or disagreement about rating or pay – 8
- The annual review should not be the first time someone finds out there are issues - 6
- try to make it developmental, growth, mentoring, coaching - 4
- Feel good if the evaluation is positive - 3
- like to provide feedback in the moment, not batch till annual review, feels wasteful - 2
- exhausted, because of tie to pay there is a lot of responsibility - 2
- depends on the recipient, some just want the rating, others want more about feedback & expectations & pay, and some are interested in alignment, direction, goals - 2
- resent the annual process - 1
- feel proud of the people I work with - 1
- not uncomfortable - 1
- view it as a learning situation for me - 1



Chi-Square Analysis of the Findings:

The purpose of the Chi-square test was to determine if a statistical significant difference exists between five various staff functions in the organizations regarding the value of the annual performance evaluation system. The five staff levels were defined as:

- Executive Management
- Human Resources
- Middle Management
- Quality Professional
- Front line staff

The researchers' perspective was influenced by the teachings of Dr. W. Edwards Deming. It was reasoned that through hypothesis testing a clear statistical difference could be demonstrated between the different

staff functions; specifically, since Executive Management and Human Resources implemented and govern the current system, it was predicted they would realize more value than would Middle Management, Quality Professionals, and the Front Line Staff. However, the test results did not prove the hypothesis. The respondent results showed a portion as expected but overall a “cross-hatched effect” was evident with some Executives and Human Resources clearly in opposition to the performance appraisal system. Furthermore, a portion of Middle Managers were in favor. None of the Quality Professionals were fans of the system.

The Survey attempted to answer - Does the annual employee performance rating and ranking system promote a culture of quality?

The culture of quality was defined as having the following 4 elements:

- Strategy Alignment
- Staff Cooperation
- Process Improvement
- Customer Satisfaction Behaviors

The respondents were asked if the performance appraisal system promoted each of the 4 culture of quality elements. The qualitative responses were interpreted by the researchers and coded as either Yes (1) or No (0). The results were blinded so no organization or employee responses were identified.

Although the total number of respondents was small (38) a Chi-Square Test of Association was used for analysis and the following 3 test criteria were met:

- Independent observations
- Categorical data
- 2 partitions were used

The Chi-Square Table below summarizes the survey results. The staff function column percentages are the ratio of respondents that answered yes to the survey questions out of the total number of respondents. The overall % of respondents' column is a cumulative of the individual percentages.

The p-value column represents the range of values for each of the 18 Chi-square test results. It was hypothesized that Executive management and Human Resources would be statistical outliers by valuing the current system more than other staff within the organization. However the chi-square test showed no staff function outliers.

The researchers were surprised by learning of Executives opposed to the annual performance rating and ranking system. With this knowledge the researchers conclude that the chi-square results do align with Dr. Deming's teachings. Regardless of the staff function there are people at all levels believing the current performance system does not promote a culture of quality.

Chi-Square Summary Table

Culture of Quality Survey Elements	Middle Mgt	Executive Mgt	Quality Professional	Human Resources	Front line staff	Overall % of respondents agree the appraisal supports Quality elements	p-value ranges
Customer Satisfaction	.13	.21	.05	.05	*	.47	.15 - .85
Process Improvement	.10	.08	0	.02	0	.21	.08 - .9
Cooperation	.05	.10	0	*	0	.21	.1 - .9
Strategy	.18	.18	.08	.08	.05	.57	.1 - .4

6 Performance Evaluations: How Is This Still a Thing? Our thoughts and conclusions on the current state and trends.

Evaluating the performance of individuals (no matter what the frequency) is not inevitable, it is a choice. Benjamin Franklin is quoted to have said, “All mankind is divided into three classes: those that are immovable, those that are movable, and those that move.”¹³ Our study identified three similar groups of organizations.

The “Movers” We identified 2 of 25 organizations that have abolished the annual process. One has substituted a system of daily management and frequent check-points (leader standard work) which includes conversations of coaching and alignment. The focus has shifted from focus on the individual to focus on the system and continuous development of people. The other organization is in the process of designing a similar focus and systems that will support that focus.

The “Movable” Several organizations of the other 23 organizations are evolving to a model of more frequent check-in conversations with a focus on continuous improvement of systems through the development of people.

The “Immovable” Some organizations seem to remain steadfast in their use of an annual performance evaluation system which has the primary focus on the individual.

There are several factors that cause the “movable” to eventually “move” and might also help some of the “immovable” to become “movable”:

- 1) The philosophical and principle-based arguments of Dr. Deming and other thought leaders. Some realize that the way that they think determines what they do, and that these mental models can be changed.

¹³ <http://www.brainyquote.com/quotes/quotes/b/benjaminfr109063.html>

- 2) The impracticalities of sustaining the prevailing systems. Some realize the waste associated with the prevailing practices, and they do what they can to remove the waste and the destructive outcomes.
- 3) The “proof” that comes from seeing that others have moved and that performance evaluation does not need to be an inevitable fixture in their organization.

There are several factors that keep the prevailing way of thinking and acting in place, that keep people “immovable.” These factors help to explain why performance evaluation is still a thing:

1. Inevitability - the belief and resulting system has a life of its own - "we've always done it this way."
2. Control and power - the belief and resulting system has an element of control and power built into it like a parent/child relationship
3. No awareness of or appreciation for systems - the belief and resulting system focuses exclusively on the individual. There is much talk about "achieving goals," and "aligning of goals." Yet, if a system was in place and was “in control”, goals discussion would not be an issue to address through a focus on individuals, but rather in system improvement. Deming said, "Changing what someone does won't change the system, but changing the system will change what someone does."¹⁴ The prevailing belief and resulting system uses performance evaluation to change people not systems, the effects of which are immeasurable and destructive.
4. Management by results - the prevailing belief and resulting system has a singular focus on measurable results and goals. There is little understanding that the most important things are unknown and unknowable, yet they must be measured.^{2,3}
5. Win-lose, competition - the prevailing belief and resulting system is based on a “zero sum” view of the world. There can only be so many “As.” The focus is on ranking and rating the individuals, rewarding those at the top, and punishing (or letting go) those at the bottom. Competition fosters silos, it doesn't eliminate them. Several survey participants cited a “calibration portion of the evaluation process” which increases conflict between individuals and departments by comparing people against others.
6. Assumptions about objectivity - the prevailing belief and resulting system hinges on a faulty claim that the process of evaluation is objective, precise and reliable.
7. Pay as a motivator - the prevailing belief and resulting system assumes that people are motivated by more pay, and that improvement will come from tying pay to performance.

¹⁴ Deming 4-Day Seminar, April 1993.