



RAISING STANDARDS • OPTIMISING ORGANISATIONS

Briefing Paper

It's a Lottery: Are you the Lucky One?



"We invented management; it's now up to us to re-invent it"
Dr W. Edwards Deming



© Copyright 2011

You may use the text in this document, which may be reproduced free of charge in any format or medium providing that it is reproduced accurately and not used in a misleading context. The material must be acknowledged as Copyright and the title of the document specified. Any enquiries relating to the copyright in this document should be addressed to Status Management Services Ltd, Harlequin House, Bickley Crescent, Bromley, BR1 2DW. Tel: 020 8460 3345 Fax: 020 3417 4897 or e-mail: cproto@stadius.uk.com

Contents	Page
About the author	2
From the author	2
Introduction	3
The current mantra	4
How do we usually view appraisals?	4
The Red Beads Experiment	5
The Results	8
Lessons Learnt	9
How performance rating skews behaviour	10
Conclusion; can appraisals ever be fair?	12
Taking action	13
Benefits	13
References	14
Additional Resources	15



About the author

Mark Woods is an apprentice served, degree qualified, engineer with masters degrees in engineering and management. Mark is a consultant and trainer who specialises in raising standards and optimising organisations, using a range of tools, techniques and standards, the core of which is systems thinking.

Mark has presented on a variety of subjects throughout the UK and internationally and is a regular contributor to a range of publications, usually contributing under the banner of "Management Matters".

He is founder and managing director of Stadius Management Services Limited.

From the author

Thank you for downloading this free briefing. This series of briefing papers represents the culmination of consultancy findings and research effort. Our hope is to educate and inform so that you can become both familiar and comfortable with ideas that may be new to you or simply to re-acquaint yourself with forgotten ideas. After reading, you will have, hopefully, set up strong foundations from which you will be in a position to move forward with your aims and ambitions.

Finally, I would greatly appreciate any feedback to mwoods@stadius.uk.com

Introduction

Since the invention of management over a century ago, management has become detached from both the day to day operation of the organisation and from delivering value to the customers who pay for it.

Conventional wisdom is that managers set targets and then create systems to monitor, measure and control the execution of these targets. These systems include budgets, performance management, incentives and appraisals, which are used to exercise control and ensure that targets are met. Simple, obvious and wrong!

We need a change in management thinking.

The current mantra

The business press advises that appraisals and performance ranking are good for business and, as a result, the practice is rife throughout business, government and educational establishments, but, just because an idea has common currency, it does not make it right.

Appraisals and rating people in particular, are a waste of time, effort and money.

Lucy Kellaway recently wrote in the Financial Times “My suggestion is dead easy and dirt cheap: get rid of the whole thing and replace it with nothing at all”.

How do we usually view appraisals?

Appraisals and rating people in particular, are a waste of time, effort and money.

Unbelievably, in some particularly hard-nosed management cultures, appraisal systems have been set up where everybody is ranked every year and the bottom X% are fired, a new tranche of people is recruited in their place and the whole sorry episode repeats the year after.

Obviously, an extreme, but what sort of atmosphere and culture do organisations that, often inadvertently, adopt this sort of approach develop?

Additionally, regardless of the atmosphere or culture, do any of us really enjoy the annual grind of appraisals? And do they really serve any purpose? Anecdotally, aren't the sentiments most of us hold about appraisals articulated something like:

- “When can we get back to the real work?”
- “I actually hate “playing God” with my people”
- “I really don't think that s/he deserves a better rating than me”
- “It's really all down to favouritism and who is the boss's pet”

Lucy Kellaway again; “But never have I learnt anything about myself as a result. I have never been set any target that I subsequently hit. Instead I always feel as if I am playing a particularly dismal game of charades, with three disadvantages over the traditional parlour game. There is no dressing-up box; there is no correct answer to guess and it isn't remotely fun. The norm is a harrowing hour's conversation during which you are forced to swallow an indigestible mix of praise and criticism referring to long-ago events, which leaves you demotivated and confused on the most basic question: am I doing a good job? The resulting form is then put on file, making you feel vaguely paranoid, even though you know from experience how much attention will be subsequently paid to it: none whatsoever.”.

The red beads experiment

In order to tease out some of the issues, let's try a little thought experiment.

Let's recruit six willing workers and have them perform work at the White Bead Company. The White Bead Company, as its name suggests, make white beads.

The main instrument of production is a wooden paddle complete with 50 depressions so that one bead will rest in each of the depressions. The production process itself requires each of the willing workers to insert the wooden paddle into a box of 4000 mixed red and white beads, the objective being to produce white beads. 80% of the beads are white and these are the beads we need to produce, however, 20% of the beads are red. These are defective beads. After each production run all beads are replaced and the box mixed vigorously before the next production run.

Production is overseen by the supervisor - let's call him Terry – and, prior to starting work, all six willing workers are fully trained in the production process. Additionally, in order to ensure consistency of production, the procedures that the workers were trained in are both standardised and documented to ensure that all staff are fully aware and familiar with their responsibilities.

Prior to starting work on the first week, Terry advises the willing workers that the management has set a target of three red beads per person per week.

The table shows the performance of our willing workers with names running down the page and the number of red, defective, beads produced, running across the page by week.

	Weeks				Total
	Wk 1	Wk 2	Wk 3	Wk 4	
Ish	16	10	7	6	39
John	9	11	12	13	45
Debbie	4	9	13	11	37
Reg	7	11	14	11	43
George	9	17	9	10	45
Gina	9	7	12	7	35
Weekly totals	54	65	67	58	
Grand total					244

Let's now look at a narrative as to how Terry, his senior managers and our willing workers may well have interpreted these results.

Week one

Obviously Terry is bitterly disappointed at the total of 54 red beads. Debbie, however, is immediately identified as someone with a special talent and is singled out as “worker of the week” and presented with a certificate. Obviously, she feels great about her new job and congratulates herself that she is settling in nicely.

Ish, on the other hand, is taken aside by Terry. He is asked “How could it be possible for you to make four times as many red beads as Debbie?” Ish is accused of being idle and work-shy at best, careless at worst. He is advised, in no uncertain terms, that his performance needs to improve. Ish himself is disappointed and feels he has let the side down. He promises himself that he’ll try harder tomorrow.

Terry points out that “training has been given, rigidly defined procedures have been written and agreed so, surely, any subsequent variation in performance must be the fault of the individual; if Debbie can produce only 4 red beads, surely everyone can”.

Week two

Understandably, Terry is extremely irritated that everybody’s performance, other than that of Ish and Gina, has deteriorated. Debbie, last week’s worker of the week, obviously comes in for most flak having rested completely on her laurels. This week she produced over twice as many red beads. Terry thinks “praise has obviously gone to her head... what was the point of the worker of the week award?”. Ish, however, has obviously taken his reprimand to heart and delivered a massive improvement producing only 10 red beads. He feels much better.

Hang on though - George has produced a massive 17 red beads! “Stop production!” Terry exclaims. A meeting is called and an investigation ordered.

The total amount of red beads produced has increased from 54 to 65. Terry points out that the expectation was that, as people became more familiar with the process, the performance would improve. Terry reminds them that “the target is three red beads each, 18 for the week. Their current performance will put the company out of business! Nobody has reached the goal. They must all try harder”.

Week three

Reprimanded and chastened once again, the willing workers return to work. There is some cause for celebration. Astonishingly, Ish improves further getting 7 red beads (obviously still more than twice the target). George also returns to form by delivering his first week’s performance of 9 red beads. However, everybody else gets worse and the total number of red beads increases to 67.

A despondent Terry informs the team that, unless performance

improves very dramatically, the facility will have to close. In order to drive up performance, he announces that he is to instigate a rigorous performance appraisal system and that people's future employment will be dependent on their performance.

Week four

Ish gets off to an excellent start on the fourth week producing only 6 red beads. Terry looks very pleased with the appraisal strategy: performance has improved dramatically as previously nearly 70 red beads were produced and now there are less than 60 beads.

However, Terry reminds the workers that the target is three red beads and that, whilst there has been an improvement, "the current situation is untenable and drastic changes have to be made". Reviewing performance as part of the new appraisal regime, he decides to fire the three worst performers. Three workers have produced less than 40 red beads and three workers have produced more than 40. John, who actually got worse week on week, has produced 45 red beads in total, George has also produced 45 red beads and Reg produced 43. He retains the three best performers: Ish, who has actually improved week on week, Debbie and Gina. The three retained workers are then put on double time.

Do you think that solution will work?

Week five

The performance table from the previous weeks has been updated to include the results from the double shifts of week five.

	Weeks				Total	Week 5	
	Wk 1	Wk 2	Wk 3	Wk 4		Sft 1	Sft 2
Ish	16	10	7	6	39	15	10
John	9	11	12	13	45		
Debbie	4	9	13	11	37	4	9
Reg	7	11	14	11	43		
George	9	17	9	10	45		
Gina	9	7	12	7	35	9	7
Weekly totals	54	65	67	58		28	26
Grand total					244		

The fifth week ends as the first week started with the production of 54 red beads (28 + 26). Terry has very mixed emotions; whilst this is the best result yet, equal to week one, it is still miles from the target of 3 red beads per person! Terry takes the results to the management and argues the case. However, he returns distraught. The management's decision to implement a performance appraisal system and then to select the best workers appears to be on the rocks. The results are not good enough and the decision has been made to close the plant.

The results

Whilst the above is, obviously, very contrived, it is highly likely that processes in your organisations, and the real world generally, will exhibit significantly more variation than the red bead experiment.

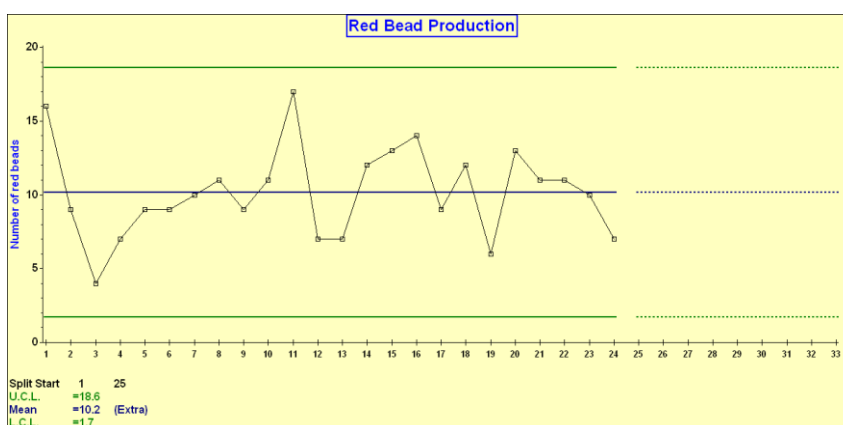
In actual fact, the red beads experiment is what would be considered a “stable” process; there are, in fact, no “special” results. Most engineers and accountants, when asked, would put the average number of red beads per week’s production at 10 (20% of 50). That would be true if the process were truly random. Oddly it’s not; it’s actually a mechanical process. As a result, when we undertake this experiment in the classroom, different paddles actually produce different long-term averages.

In this instance, for the first four weeks, the average of this stable process is 10.2 (the total number of red beads divided by the number of man weeks worked (244 divided by 24 (4 Weeks x 6 people))).

....past data can be used to predict future performance.

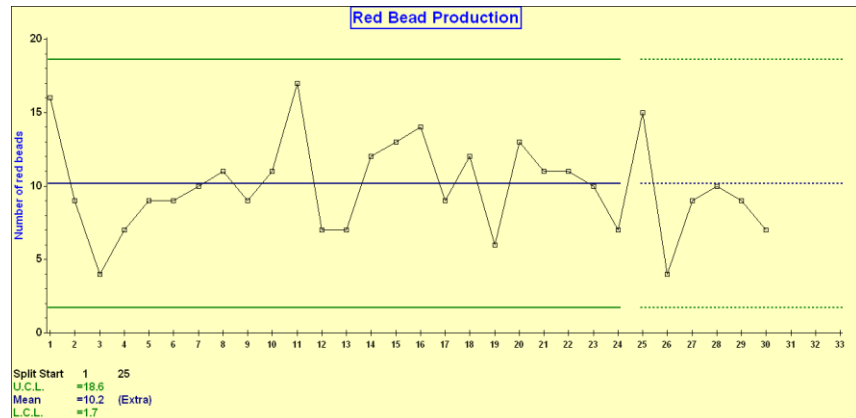
Again, taking the results from the first four weeks, it is also possible to calculate from the data the limits at which a result either higher or lower would be deemed to be exceptional. You might be relieved to know that I am not going to do the maths here: if you do want to know more about the calculation, please call. Otherwise take it from me that, with these results, the variation ranges from 1.7 at the lower end to 18.6 at the higher end. That is, any result between these figures is entirely reasonable and to be expected. None of the results from any of our willing workers is out of the ordinary or exceptional.

A graph, whose technical name is a control chart but which we prefer to call a process prediction chart™ can be seen below. It shows the individual results, the average and the upper and lower limit (between which all results are “normal”).



Indeed, the results from weeks 1-4 can be projected forward, as seen by the dotted lines, so that past data can be used to *predict* future performance. We can now predict that, if nothing changes, this process will continue to deliver between 1.7 and 18.6 (2 and 18) red beads.

The data from the double shift of week five has now been added into the chart and we can see that the variation is as expected.



As management, we might not like the range of variation and that's the message left by Dr Deming; it's the job of management to reduce variation (and move the average up if up is good) and, after making changes, the process prediction chart™ can show us if we have succeeded.

Lessons learnt

There are a number of lessons to be learnt from this simple thought experiment.

In this instance, and many others, genuine improvement will only result if there is a genuine change to the process itself.

...people are compared to one another when, in actual fact, they are helpless to affect their performance. The results are entirely dictated by the system within which they work.

- Not one of the workers had any more chance to do any better than any of the other workers.
- It is critically important to appreciate just how big the variation can be in a stable process. It is reasonable to occasionally be as bad as 17-18 red beads. Conversely, you don't have to be especially gifted to get as few as 2 or 3 red beads.
- Most processes in the real world exhibit far more variation than that found with the red beads experiment, so swings in variation in the real world will be even more dramatic.
- We are all very well practiced in making up descriptions for figures that are often the result of stable processes. The outcome is that, whatever the result, we find reasons to explain the figures when, in fact, there are no reasons.
- Most times there is a 50/50 chance that the next result will either be higher or lower than the previous one.
- There is (usually) no point in feeling "good" or "bad" about a particular result if it is within the bounds of expectation.
- Despite his best intentions, there really was no point in any of the protestations or congratulations made by the supervisor, Terry.
- The investigation into George's result of 17 red beads was pointless as the result was within the bounds of expectation.
- It was completely unjustified to fire the workers for their "poor" performance – their performance was beyond their control.
- The cause of the problems is the number of red beads in the system. Work needs to be done to get the red beads out of the system.

Time and time again people are compared to one another when, in actual fact, they are helpless to affect their performance. The results are entirely dictated by the system within which they work.

The time and effort spent commenting on results of this nature without properly understanding the range of variation is entirely wasted. The same time would be more profitably employed working with staff (and suppliers) to remove the red beads and improve the system.

In this instance, and many others, genuine improvement will only result if there is a genuine change to the process itself.

How performance ratings skew behaviour

The idea of performance appraisal is alluring. But, just because an idea has common currency, does not make it right.

- People pat themselves on the back for performances that are not special
- Conversely, people feel bad and blame themselves and others for performances that are actually not so bad
- People rated below average take a look at others around them and wonder why the difference exists

...it results in the abdication of the real responsibility – to improve the system.

Rating people is often based on a “count” of some kind often with a comparison, intentional or otherwise, to targets or to other people, for instance:

- The number of calls made
- The value of sales made
- The number of widgets produced
- The number of drawings completed
- The number of service visits made

Any index that results in people being measured by some kind of “count” eliminates the chance of people taking pride in their work. There is no chance of review, no chance of checking, as, to do so, would decrease the workrate and therefore performance. Counts are actually easy, but result in the abdication of the real responsibility – to improve the system.

An allied point, as recently noted by John Seddon, is “if we design work into functions and give each function its own target, should we be surprised if they don’t co-operate with one another”? This is as true for functions, departments and individuals. Rating one another destroys teamwork. If your rating is geared to your performance, what incentive is there to work with others in the organisation to improve processes?

Top level jobs, and the associated top level salaries, are the (understandable) goal for many of us so, given that your entire future depends on getting a good rating from the boss, your entire focus becomes getting that good rating so that you can progress through the organisation’s rankings. People may also become afraid to challenge the boss because to challenge the boss may affect their rating and therefore their promotion. It’s subtle, but very significant; the aspiration and the emphasis are placed on getting a good rating to facilitate further promotion, not to improve the way in which the organisation works. It’s safer not to rock the boat and work for your own promotion. The down side is that the organisation becomes very risk averse.

There is no attempt to improve the system nor is there any attempt to reward people for doing so.

Conclusion: Can appraisals ever be fair?

At Status we are not as strident as Lucy Kellaway about abolishing the annual job chat, but we are certainly not in favour of rating people without understanding the system within which they work. The first question we would ask is “what is the objective of the review?” To assess learning, to assess satisfaction, to assess desire for development are all legitimate aims, but the system needs to be designed accordingly.

The red beads experiment is ridiculously simple, but it makes the anti rating point stunningly well; once people appreciate it, they find red beads, over which individuals have no control, all over their organisations.

Paul Moore, the HBOS “Whistle Blower”, reported on “The Choice”, Michael Buerk’s Radio 4 programme, the practice of “Cash or Cabbages”. Presentations were made where those that had met their performance targets got cash whilst those that had not were awarded cabbages. Mr Moore spoke graphically about the culture of fear that developed - performance appraisals nourish fear and create politics.

Additionally, as was so elegantly argued by the Hopper brothers in their excellent book the Puritan Gift: Reclaiming the American Dream amidst Global Financial Chaos, “In any well run organisation, an individual’s achievements were likely to be due as much to the wisdom with which he was directed from above, and to the support of his equals and subordinates, as to his own efforts.”.

“there is a mismatch between what science knows and business does”

Dan Pink

As Dan Pink, the freelance thinker and devil’s advocate suggests, with a characteristically pithy one liner, “there is a mismatch between what science knows and business does”. Appraisals don’t work: there is too much variation in the system.

Finally, one of the most famous people thinkers, Fredrick Herzberg, undertook a number of studies on payment for performance and found that it runs counter to our nature:

“It is only when he has his own generator that we can talk about motivation. Then he needs no outside stimulation: he wants to do it.”

If you want further “proof”, grab a cup of coffee and take a look at the following:

- Youtube.com "Honk if you love performance appraisals" by Ken Miller.
- <http://www.youtube.com/watch?v=m25LrJAH1D8>
- TED.com "The surprising science of motivation" by Dan Pink.
- <http://www.ted.com/search?q=The+surprising+science+of+motivation>

Taking action

To paraphrase an observation from Deming: often, management know everything about their organisation except how to improve it.

The aim of leadership is to improve performance, to improve quality, to increase output and to bring a pride of workmanship to people which will further improve performance.

In order to adopt an approach of this nature, a number of steps are likely to be required, for instance:

- The top team needs to establish the purpose of the organisation, that is the *benefits* and *capabilities* that the organisation is looking to deliver to the customer; understood from the customer's point of view
- The purpose needs to be checked with a select number of friendly customers
- The key measure, or measures, which relate to the purpose need to be established
- Sub measures may need to be established by working backwards into the organisation
- Measures need to be related to key processes
- All measures need to be seen on process prediction charts™
- Teams (of the people undertaking the work) need to be established in order to improve processes
- The results need to be monitored and success celebrated

Benefits

Status works with client companies and organisations to embed a culture of performance measurement and management that delivers real results in a manner that leads to joy in work and joy in learning.

Implementing a robust performance measurement and management can deliver staggering results.

- Customer and stakeholder satisfaction significantly improve
- Staff motivation and productivity improve as people are allowed to:
 - solve problems
 - improve the way in which their work works
- Profits and productivity are raised

Essentially, the focus shifts from “making the numbers” to “meeting the purpose, learning and improving”.

How much has this paper whetted your appetite for fairer people processes? Engage with us and become part of the story! For a consultative meeting or additional information, please contact Mark Woods on 07976 426 286 or email him at mwoods@stadius.uk.com.

References:

Deming, W. Edwards. *Out of the Crisis*. Cambridge MA: MIT Centre for Advance Engineering Study, 1986

Herzberg, Fredrick. *One more time. How do you motivate employees?* Harvard Business Review (January-February 1968) pages 53-62

Hopper, Kenneth and Hopper, William. *The Puritan Gift. Reclaiming the American dream amidst the Global Financial Chaos*. I.B. Tauris, 2007 page 135

Joiner, Brian L. *Fourth Generation Management*. McGraw Hill Inc, 1994

Kellaway, Lucy. *It's Time to Sack Job Appraisals*. Financial Times July 11, 2010

Neave, Henry. *The Red Beads Experiment*. Deming Forum publication A15 (Revised 2009)

Seddon, John. *Back to the Method*. Quality World, Sept 2009

Performance Appraisal and all that. The British Deming Association publication A8 (Revised 1992)

Additional Resources:

Are you in chaos, clarity or confusion?
Review your organisation's performance;
take 10 minutes with the Status Coffee
Break Challenge at:



www.status.co.uk/coffeebreakchallenge

The Coffee Break Challenge is a questionnaire which been designed to provoke thinking about your organisation's current performance. Be honest with yourself. It is deliberate that there are no scores: the challenge is designed to make you think. There is no one looking or checking!

Additional Briefings:

It's broken – Housing repairs and other field service operations

An examination of systems thinking as applied to housing repairs and other field service operations. In the housing arena a plethora of Government targets is actually hampering the effort to improve. This paper seeks to return to basics, that is, to define the "purpose" of the system and, from there, create management systems that deliver value to the tenant or client.

David and Goliath: Optimisation 3D™ and Six Sigma

Six Sigma has mixed reviews in the press. This paper seeks to examine the fundamental focus of Six Sigma and contrast it with the Status process, Optimisation 3D™, whose focus is to delight the customer.

Creating competitive and compassionate contact centres

Contact centres play a critical role in many firms and sectors. However, they are often labelled as the "sweatshops" of modern business industries offering repetitive, pressured and boring roles with little, if any, career progression. This paper applies systems thinking to contact centres in order to create competitive *and* compassionate environments.

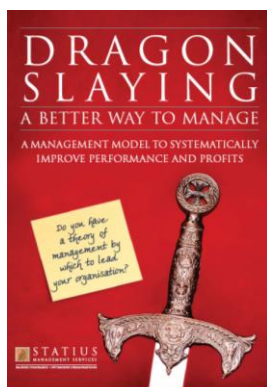
Targets, goals and other management myths

Conventional wisdom is that managers set targets and then create systems to monitor, measure and control their execution. These include budgets, performance management, incentives and appraisals, which are used to exercise control and ensure that targets are met. Simple, obvious and wrong! This paper sets out a systems thinking alternative.

Easy Meat? Cutting the Fat in Construction

The purpose of Lean Construction is to increase capacity by designing the construction process to optimally respond to customer demand. So, if an organisation can cut even just small chunks from the 55-65% of work that the Lean Construction Institute estimate is used to produce waste, staggering results can be obtained. This paper explores that debate.

Dragon Slaying



Dragon Slaying is Mark's long-awaited book which picks apart a number of management myths. The benefits in adopting the ideas in the book are:

- A more informed understanding of how an organisation delivers value to customers and stakeholders; how the work in an organisation works
- The development of a strategy for "Listening to Customers and Stakeholders"
- The development of the organisation in which everyone's efforts result in:
 - Improved performance
 - Less stress
 - Improved profit