Making what's Counted Count

Ian Graham MIEE, CEng, Visiting Fellow at the Centre for Business Performance of Cranfield University, Principal Consultant with PRISM Consultancy Ltd.

"Not everything that counts can be counted and not everything that can be counted counts" – Albert Einstein (in a frame on his wall at Princeton)

Synopsis

One of the consequences of the introduction of the Balanced Business Scorecard has been a growing interest in the development of wider and more sophisticated measurements in organisations in order to enhance performance and focus ever reducing resources. The Centre for Business Performance at Cranfield University has made a major contribution in this field providing considerable information on the types of measurement that can be considered. However, knowing what data to collect is one thing; knowing how to interpret it is another. What is missing from many managers' competence (and this includes managers at all levels up to the top) is knowledge of the Theory of Variation. Without this knowledge many managers can and do interpret data incorrectly – often and unknowingly using it as the basis for incorrect action. In addition, they do not have access to the powerful new levels of interpretation which knowledge of this theory brings. This paper sets out to explain in simple terms what this theory is and how it contrasts with the implicit theory we use today.

Two Theories Regarding Numbers in the Real World

There are two currently held theories regarding the use and interpretation of data emanating from human activity. Actually, those that follow the first of these theories are unaware that there is a theory at all – they just think they are using common sense when it comes to looking at data in the real world. Those that understand the second theory realise there is a different way of interpreting data in the real world and, because they understand the second theory, become extremely uncomfortable with those who continue to use data in the old way. We need to explain the two theories in simple terms and show how a practitioner may move away from Theory 1. In this way, we can leave behind what are ineffective and destructive behaviours connected with this theory and move towards Theory 2. This will introduce far more powerful ways of using data to drive performance.

Theory 1 – It's all arithmetic

Everyone is collecting data – and the advent of the computer has made it possible to collect data in lorry loads and analyse it to death. It is worth noting in passing, and with reference to Einstein's quote in the heading, that often it is the wrong data that is being examined in the first place. However, for the purpose of the discussion here, we shall assume that the data being considered is useful as a basis for understanding performance and for action and improvement.
Fundamentally, Theory 1 says that every single point of data is specific and means something.

Theory 1: Every single point of data is specific and means something

This sits happily with the vast majority of people who didn't get much further than arithmetic at school. Such people do, however, understand the actual precision of numbers and expect them to behave precisely in the real world. They well remember from their school lessons the richness of words because of all the shades of meaning they contain whereas they also recall that numbers are boringly precise and consistent wherever they are encountered.

Simple arithmetical computation such as calculating percentages is fine and consistent with this theory. So is expecting not only that every number has a single reason for its existence but that the difference between two numbers will also have a single reason. This argument gives rise to business analysis and action based on comparisons between two numbers and also making the comparison based on arbitrary standards and timescales. Examples of these comparisons include:

- This month compared to last month
- This month compared to the same month last year
- This month compared to a target or budget

These comparisons are often reinforced by looking at the same things on a year-to-date basis. By arbitrary timescales we mean that it is common to make the comparisons relate to the position of the Earth in its orbit round the sun or the Moon around the Earth namely: monthly, quarterly, annually and so on. The Earth and Moon don't actually know where they are in their orbits but users of Theory 1 apparently think these positions are significant. Arbitrary standards are the targets that are set that are little more than guesses or negotiation outcomes.

For an illustration of this thinking, consider the table below. All managers will be familiar with this reporting format and one can often observe besuited individuals on trains and planes poring over such data reports. They will mark them with highlighters and make notes in the margin and one can tell they are raising questions or preparing their defences for the next management meeting. I have shown two years of data so that the points raised make sense especially when we move to Theory 2. This is actually unusual – managers live in a one-year snapshot and the only reference that is made to historical data is the percent comparisons year-on-year.

The data are shown as being a record of “Defects” but it could be information to do with anything that the operation finds undesirable, such as:

- Rejects on the production line
- Late payments
- Accidents or injuries in the plant
- Overdue invoices
- Deaths on the roads
- Typographical errors in reports
- Errors in design documents
- Incorrect deliveries
- Etc, etc……..

In this scenario, therefore, up is bad and lower results are a cause for celebration.

Month	Defects (100s)	Monthly target	Monthly % variance from target	Month % diff from last year	YTD	YTD target	YTD % variance from target	YTD % diff from last year
Jan-03	28	28	0.0		28	28	0.0	
Feb-03	26	28	-7.1		54	56	-3.6	
Mar-03	28	28	0.0		82	84	-2.4	
Apr-03	26	28	-7.1		108	112	-3.6	
May-03	30	28	7.1		138	140	-1.4	
Jun-03	30	28	7.1		168	168	0.0	
Jul-03	32	28	14.3		200	196	2.0	
Aug-03	30	28	7.1		230	224	2.7	
Sep-03	30	28	7.1		260	252	3.2	
Oct-03	29	28	3.6		289	280	3.2	
Nov-03	32	28	14.3		321	308	4.2	
Dec-03	27	28	-3.6		348	336	3.6	
Jan-04	31	26	19.2	10.7	31	26	19.2	10.7
Feb-04	27	26	3.8	3.8	58	52	11.5	7.4
Mar-04	35	26	34.6	25.0	93	78	19.2	13.4
Apr-04	24	26	-7.7	-7.7	117	104	12.5	8.3
May-04	25	26	-3.8	-16.7	142	130	9.2	2.9
Jun-04	28	26	7.7	-6.7	170	156	9.0	1.2
Jul-04	31	26	19.2	-3.1	201	182	10.4	0.5
Aug-04	26	26	0.0	-13.3	227	208	9.1	-1.3
Sep-04	27	26	3.8	-10.0	254	234	8.5	-2.3
Oct-04	29	26	11.5	0.0	283	260	8.8	-2.1
Nov-04	27	26	3.8	-15.6	310	286	8.4	-3.4
Dec-04	30	26	15.4	11.1	340	312	9.0	-2.3

Figure 1 - A typical management data summary

Notice the arrival of a “stretch” target at the beginning of the second year. This may well have been based on the apparent evidence that a performance of 2600 defects per month is “achievable” from the data of the first four months of the previous year.

The kind of questions and issues that are raised by the consideration of this type of report will occur not by looking back at the shape of results over the year but will happen month by month as the latest result is announced. Remember, every result means something by Theory 1. Such a monthly commentary might look like this:

- January was very poor – well above target. This was due to pressure on production before Christmas.
- February shows an improvement and is close to the February result of last year. We are optimistic this improvement trend will continue.
- March was a disaster – what happened? We will call an emergency management meeting where we will study the data and call the worst offenders to task.
- The April result clearly shows the payoff from our tough actions taken last month – 7.7% below target!
- More improvement in May – not only nearly 4% blow target but 16.7% below May last year! Everyone should receive a letter of thanks or a bonus.
- A disappointing June. We beat the target in April and May – why can't we beat it every month?

And so on, month by month. Later in the year more defensive commentary might appear as the end result comes into view, such as:

- On a year-to-date basis, almost the whole of the second half showed an improving trend. If we keep this up we're bound to show real gains in 2006.
- December is a real disappointment. The first month in nine not to show an improvement on the same month a year ago. However, the evidence of the past nine months indicates that we should expect continuing improvement as we enter 2006. This will be further encouraged by the management approved stretch target of 2400 defects per month for the new year.

The Problem of Fluctuations

People who work within the thinking of this theory are aware of irritating fluctuations in their data. They can choose to ignore them, which is not difficult because they view the data mostly in snapshot terms, or those who have an inkling that there may be more to this than looking at individual numbers will attempt to smooth the data in order to see longer term movements such as trends. Computers make this smoothing simple and any user of Excel can have easy access to trend lines, either linear or polynomial, and moving averages both linear and exponentially weighted and over any time period. This puts tremendous sophistication in the hands of the mathematically challenged which enables people to bamboozle themselves as well as others. Let's have a look at the above data on a graph constructed in Excel:

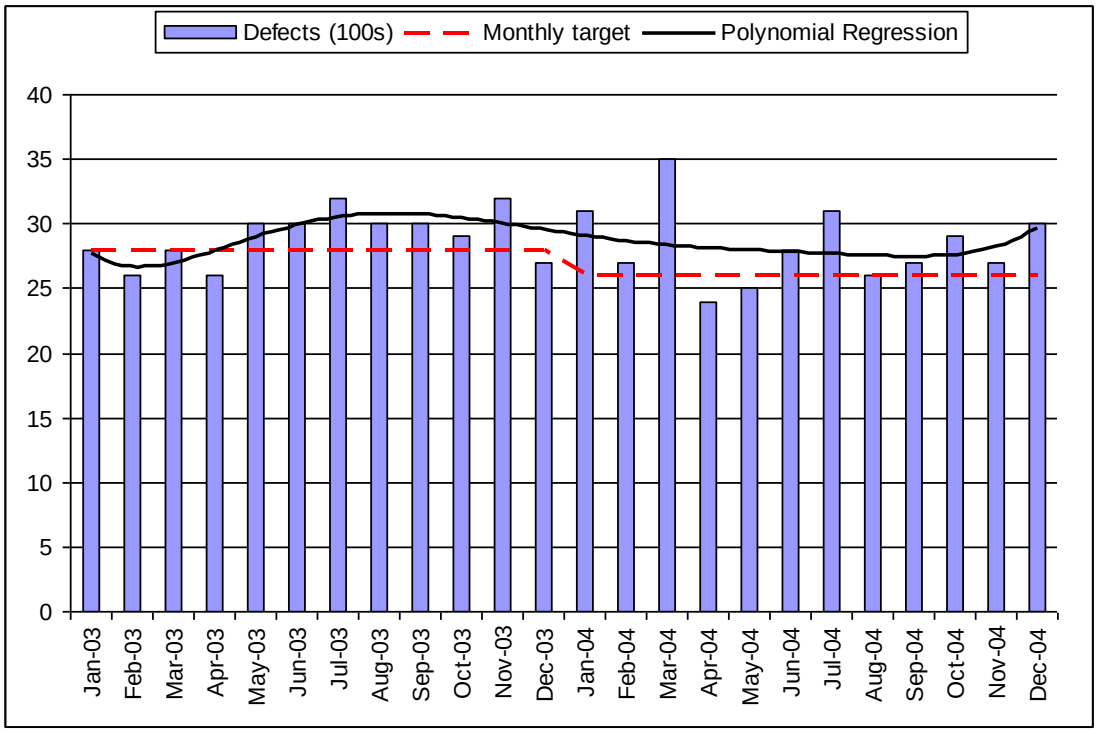

Figure 2 - Data from Fig 1 presented as a time-series graph

The managers are right to want to look at pictures of the data – but not for the reasons they imagine. By looking at the data in graphical form they are beginning to look for patterns rather than individual numbers. However, they now face a new problem (but remember they don't know they have this problem because it is still firmly entrenched in Theory 1 analysis). The problem is, because they still think that every number means something, they think this is also true of every device that Excel makes available to them: linear trend lines, lines of regression, moving averages and so on. They think they will be able to see patterns like trends by looking through the fluctuations. **The problem is they don't realise that the patterns they can see this way might well be illusory and will provide incorrect information regarding performance.**

The targets for the two years are shown on the graph, as is a line of fit – in this case a polynomial regression line (how's that for sophistication?!) The kind of questions and issues that are raised by the consideration of such a graphical analysis might be as follows:

- The regression line shows that we start the year off well but things get worse in the second half. Is there some kind of cycle or seasonality to this defects problem?
- March was a disaster – what happened? It is the highest point in the entire sequence. What does this mean? Can we just treat it as an outlier or a "blip"?
- We started off 2003 so well compared to target – why did 2004 start off so badly?

- Things have definitely been getting worse since April as the improving trend slowed down then reversed. What do we have to do to turn around this pattern?

It's worth noting that if we had used a simple Excel linear trend line over the two years it would have shown a gentle improvement – probably giving cause for great optimism!

As with the commentary on the single data points, the managers are assuming every movement of the curves and lines means something. I repeat, there is no denying the accuracy of the arithmetic – it is correct – the problem with Theory 1 is it based on the assumption that one can treat the analysis and interpretation of the real world data arithmetically also. Therein lies the flaw. One has only to listen to the TV Business News to hear the same kind of commentary: "The FTSE Index fell yesterday because of nerves over oil prices." Or, "We could be heading for recession because retail sales in April were 2% below the same month last year."

"All models are wrong, but some are useful"

Prof. George Box said, "All models are wrong but some are useful". Action based on Theory 1 is being widely applied in business, government and industry today and it would seem that people are getting by with this approach in the main. After all, for centuries people were able to travel widely assuming the world was flat. Theory 1 holds sway and nobody challenges it. But then, how could they know? It's only when one understands Theory 2 that the problems associated with Theory 1 make themselves evident. More importantly, Theory 2 holds the potential for the identification of completely different questions and dramatic new actions for real improvements in performance.

So what is this Theory 2?

Theory 2 – It's actually statistics - or rather: statistical thinking.

The clue lies with the "irritating fluctuations" mentioned above. The theory starts with observing that the data derived from real world situations is always infected with variation. It proposes that this variation is brought about both by chance combinations of a multitude of factors inherent to the system under consideration, and also by significantly different events that may be a part of the system or can hit it from outside. What this means is, and this is Theory 2: There is going to be variation in all data from the real world. However, a few single numbers are specific and mean something (the significant events – sometimes called "signals"). Other numbers mean nothing individually but mean a great deal collectively (the random combination variation inherent to the system – sometimes called "noise").

Theory 2: A few single points of data are specific and mean something. Other points of data mean nothing individually but mean a great deal collectively

Often the random component in the data is quite large which makes it hard to spot the signals. The response to the two types of variation must be different: signals, more correctly called Special Causes of Variation, are "assignable" to a specific reason. Actually, Theory 1 behaviour is fine for Special Causes – each one means something so that it can be dealt with appropriately: deliberately choosing to disregard it once it is identified, ways found to avoid it happening in future or learned from to bring about improvements. Response to noise (more correctly called Common Causes of Variation) is different. Fundamentally, all the Common Cause points of data are manifestations of the same thing and are indications of how the process from which they come is behaving over time. In other words, it is wrong, futile and even dangerous to react to individual points of Common Cause variation. What you must do instead is react to its position, namely its average, and its spread in order to drive performance towards the desired level.

To summarise, it is possible to make either or both of two mistakes:

1. Treat Special Causes as Common and either ignore them entirely or build complexity and duplication into the system in case they ever happen again.
2. Treat Common Causes as Special and treat every twist and turn of the data as meaning something and adjusting the system each time. This latter effect is called **"tampering"** and the effect of tampering is to make the variation greater!

Now, the astute reader will realise that both these two mistakes could be being made in the scenarios described above.

Once one understands Theory 2 and that all data from the world of work and business is infected with this Common Cause variation we realise that, in order to bring about improvements in our performance we must avoid these two mistakes and, instead:

1. Buy into Theory 2 and accept the existence of the variation and;
2. Learn how to tell the difference between the two different types of variation AND know what new actions to take based on this knowledge.

But now we have a new problem. If we can fall into the trap of making these mistakes we need to avoid them by being able to make the distinction between the two types of variation. But how?

Fortunately there is help at hand.

The Process Behaviour Chart

Many managers have encountered Statistical Process Control (SPC) during their careers. They've probably parked it in their minds as something one does in high

volume manufacturing when understanding the capability of machines to produce items within specification. While this is true, in recent years there has been a growing understanding of the power of this technique when dealing with data from processes everywhere (including in transactional and service environments) and also in with the data that management uses to drive the business. I would go so far as to say that all management data from the Balanced Scorecard or other Key Performance Indicators should be evaluated using this technique.

For those unfamiliar with this approach let me explain it simply. It was invented by an American called Walter Shewhart in the 1920s and was used effectively in that country for the manufacture of munitions during the Second World War. After the war, the Americans appeared to have forgotten the technique but it was picked up by Japanese manufacturers and is at the heart of their ascendancy in automotive and electronics manufacturing success up to the present day. Recently it has become of great interest in many different fields as indicated earlier. The device that is used in SPC is called the “Control Chart” – more recently re-titled the “Process Behaviour Chart” as being more descriptive of its purpose. There are a number of these charts depending on the specifics of the application and the type of data being evaluated but here we will describe the one that is appropriate for analysing management data such as in our example. This chart is called an “Individual Values” or, more technically, an “X-mR” Chart. It is used where the data are available singly and in sequence such as in monthly reports.

The chart is a simple line graph where the data is plotted as a time series. Already we see an issue of importance, which is that the chart encourages us to consider the data of the present in the context of the past and presents us with the potential for looking into the future. A simple pictogram of such a chart is shown below:

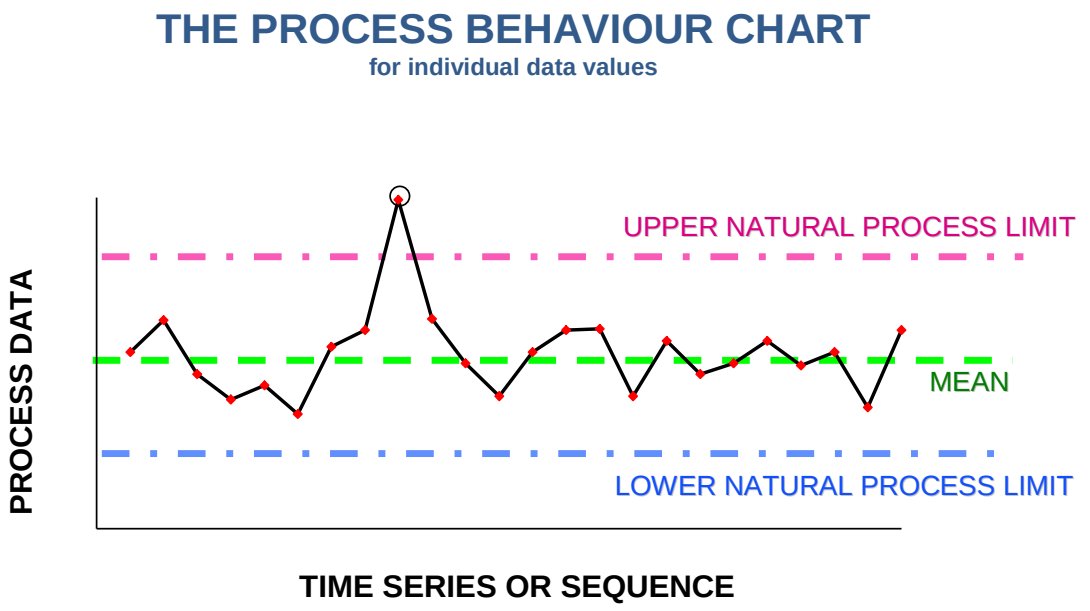

Figure 3 - A simple diagram of a Process Behaviour Chart

The graph exhibits variation. This is no surprise as it is a consequence of Theory 2. The average line is shown – normally the arithmetical mean as are two other lines called the Upper and Lower Natural Process Limits (known in the traditional SPC world as upper and lower "Control Limits"). Everyone understands the average line and how to calculate it. The two Process Limits show the natural excursion of the data due to the variation. As with the average, these limits are calculated from the data itself and they are not specification or target lines set from outside. Shewhart's guidance was to set these limits at plus and minus three Standard Deviations from the average – Standard Deviation being the measure of spread or dispersion of the data. We are not going to cover how the calculations are made in this paper because we want to concentrate at this time on how to use the charts for interpreting business data. The key point is the limits give us a definition of when to look for a Special Cause. In the figure, the Special Cause is the point outside the upper limit.

We now have a tool for distinguishing between Special and Common Causes of variation. Let us apply it to the business data in the scenarios above:

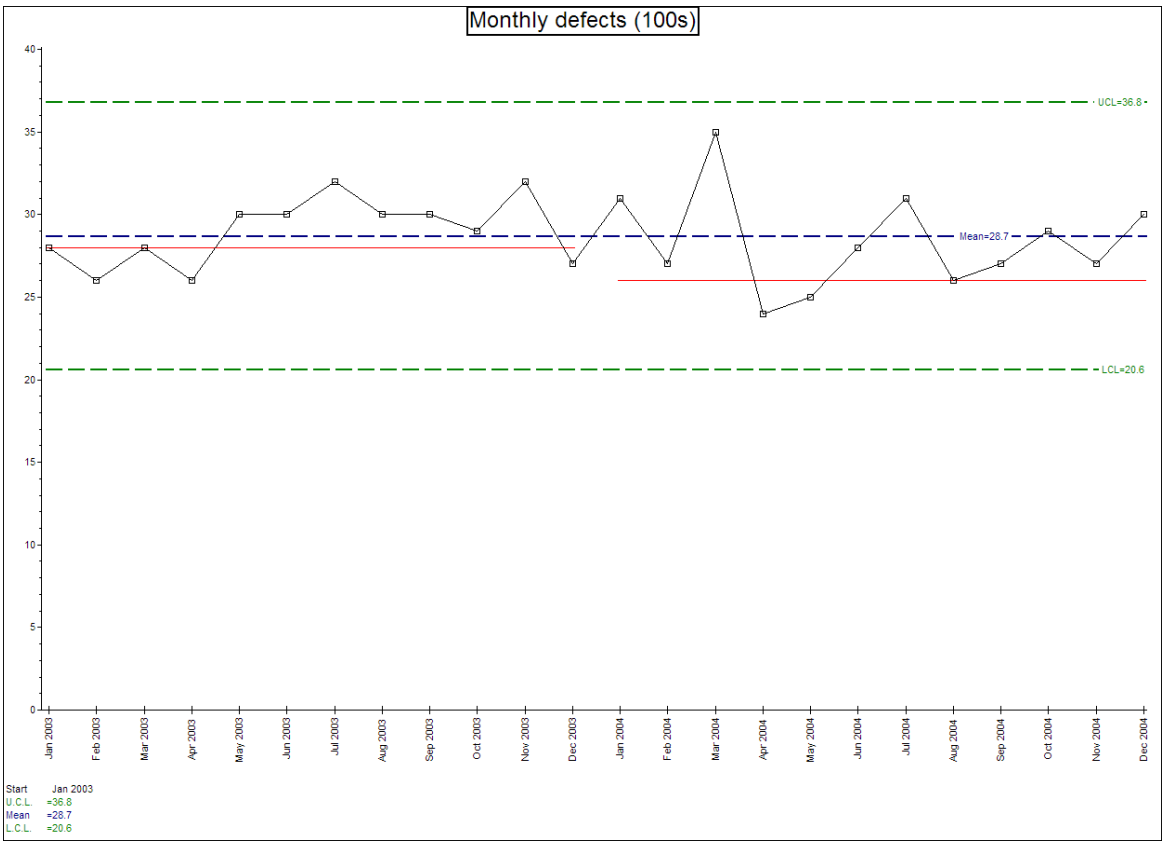

Figure 4 - A Process Behaviour Chart of the monthly defect data

Notice that we have shown the target lines in red. Also notice that target lines pay no attention to the presence of variation – they come from Theory 1!

The kind of questions and issues that are raised by the consideration of this graphical analysis might be as follows:

- The average for the system we have in place which produces defects is 2870 per month.
- We expect the defects to fluctuate between 2060 and 3680 and any result between these values will be no surprise (including the March 2004 result).
- All the variation is Common Cause and any action based on a single result such as investigations or blame will make matters worse by increasing variation.
- The target for 2003 is just below the mean. This means that the target will be failed about half of the time and beaten about half of the time. The chances of beating the “stretch” target in 2004 are even lower.
- Most importantly, unless we do something we will get more of the same – an average of 2870 defects per month; and monthly outputs that vary between 2060 and 3680 defects.

On this final point, it is worth referring back to the author of our quotation at the beginning of this paper. **Albert Einstein** also said:

"Insanity: doing the same thing over and over again and expecting different results."

In many organisations the only device that is used to stimulate and improve performance is the stretch target – the assumption here being that performance comes from peoples’ efforts and that the effect of the target will be to get them to try harder. Such an assumption leads to a cycle of despair as described earlier. The application of Theory 2 and the Process Behaviour Chart is based on the assumption that **performance comes from the system that people work in**. If this assumption is correct then the PB Charts are there to aid improvement by working on the system and introducing real change. On this basis, the discussion among the managers in our final data scenario would go something like this:

- The only target for defects that makes any sense is zero. Our customers won’t accept defects and they are costly waste in our system. We may never achieve zero but our goal must be perfection.
- The way we will move towards our goal will be by:
  - Understanding how our systems and processes can produce defects.
  - Gathering and stratifying the data for the types of defect we are producing.
  - Using systematic methods for data analysis.
  - Inviting our people to help us understand what is happening to produce these defects and listening carefully to their knowledge and ideas.
  - Proposing and introducing changes in our processes that will bring about improvement.
  - Evaluating the improvements with the aid of Process Behaviour Charts
- We must establish effective governance to hold the gains we have made and as a springboard for further improvements

Conclusion

The move towards understanding variation and the use of Process Behaviour Charts is a major change from the traditional ways of analysing data and using it to understand performance and drive improvements. It is essential that people are educated and taught the techniques described in this paper and the other methods that are so effective in improving processes within a systematic approach. Let us not fool ourselves here: just learning a new technique will not change us. It is difficult to unlearn the habits of a lifetime. Moreover, if we accept and begin to apply the thinking and practice of Theory 2, we may well be operating in an environment that is still operating with Theory 1 and insisting that we respond in the same way. In the end, there are going to be major cultural as well as practical hurdles to overcome – but that is not a good reason for not doing it.

Ian Graham
8th June 2005

References

The Performance Prism – Neely/Adams/Kennerley. Prentice Hall ISBN: 0-273-65334-2
The Balanced Scorecard – Kaplan/Norton. Harvard Business School Press ISBN: 0-87584-651-3
Understanding Variation – Donald J Wheeler. SPC Press ISBN: 0-945320-35-3
The Leaders Handbook – Peter R. Scholtes. McGraw-Hill ISBN: 0-07-058028-6
Fourth Generation Management – Brian L. Joiner. McGraw-Hill ISBN:0-07-032715-7