F

DAY 3:

UNDERSTANDING VARIATION

and

THE FUNNEL EXPERIMENT

DAY 3: UNDERSTANDING VARIATION AND THE FUNNEL EXPERIMENT

(Stats-level 0 only)

(9.30am – 1.00pm; 2.00pm – 6.00pm)

NB In the text, the clock icons for **Stats-level 0 only** *are on the left-hand side of the pages during the morning of Day 3. In the afternoon the timings for all Stats-levels are the same and so appear on the right-hand side as usual.*

DAY 3: UNDERSTANDING VARIATION AND THE FUNNEL EXPERIMENT

(Stats-level 1–3 only)

(9.00am – 1.00pm; 2.00pm – 6.00pm)

NB In the text, the clock icons for ***Stats-levels 1–3*** *are on the right-hand side of the pages as usual.*
The clock icons on the left-hand side during the morning of Day 3 are for ***Stats-level 0*** *only .*

(Stats-level 0 only) (Stats-level 1–3)

DAY 3: UNDERSTANDING VARIATION AND THE FUNNEL EXPERIMENT

Variation—the enemy of quality

I'll start by telling you of something extremely basic that puzzled me in the 1980s when I was just beginning to learn about Dr Deming's work. It was a basic matter involved with Statistics. Now, I had already been a University Lecturer in Statistics for some 12 years before ever hearing of Dr Deming. Despite that background, for a considerable time I couldn't understand what was arguably the most important of all the messages in his four-day seminars. It was the need to concentrate on *understanding variation*—apparently more than on anything else. Why was I puzzled?

As already indicated on Day 1, much of the way Statistics is usually taught involves [A] data and [B] probability and probability distributions. (Again, as I made clear at the time, if you know little or nothing about [B] then that doesn't matter as far as this course is concerned—it might even be an advantage!) In both cases I was always used to working with *averages* and with *variation*—*both* of them but with averages being the more important. So I could understand Dr Deming's interest in variation, but I couldn't understand why he hardly ever mentioned the average. Particularly when considering *processes*, of course it is good to reduce the amount that things vary, but surely it's even more important to get the average right. So why didn't he say anything about that?

Eventually, after many months, light dawned—through thinking about something from long ago: my school-days!

The school bus

Throughout my schooldays, I lived out in the countryside, several miles from my school. School started at 9.00 each morning. I travelled to school by bus. The bus would usually arrive at the bus-stop near my home some time between 8.25 and 8.35, though occasionally a little earlier or a little later. Now and again, it was *very* late, e.g. because it had broken down, or had been held up by a road accident, etc. Other than on such rare occasions, it would always get me to school on time. But I was not happy!

To be fairly sure of catching the bus, I had to be at the bus-stop by 8.25. To be *really* sure of catching it, I needed to be there by 8.20. But much of the time I'd then be waiting 10 to 15 minutes, or occasionally longer—often when the weather was cold and raining. (It seemed to rain a lot in England when I was young—and this was a bus-*stop*, not a bus-*shelter*!) So I could well be soaking wet and/or freezing cold by the time the bus arrived. Not a good start to the day!

If only the bus could have arrived within, say, *one* minute either side of 8.30, rather than within 5 or even 10 minutes. Or one minute either side of 8.25, or one minute either side of 8.35, or indeed one minute either side of *any* suitable average time of arrival. Then I could have arranged my mornings much more efficiently—and, with very rare exceptions, suffered no more than a *two*-minute soaking!

So I learned that the *variation* in the arrival-time of the bus seriously affected my quality of life! Note that the variation was actually more important than the *average* time of arrival. My process was easily adaptable to whatever the average happened to be. But the greater the variation, the more I risked either missing the bus altogether or getting wet through—*irrespective* of the average.

Variation is indeed the enemy of quality.

BACK TO THE WESTERN ELECTRIC COMPANY

Next, as promised on Day 1, here is a rather longer extract from Dr Deming's account of what happened at the Western Electric Company back in the 1920s. This is again transcribed directly from the presentation he gave in Versailles in July 1989 (see BDA Booklet A6 pages 2–3). You have seen part of it before but, to put it mildly, it is worth repeating:

> "Part of Western Electric's business involved making equipment for telephone systems. The aim was, of course, reliability: to make things alike so that people could depend on them. Western Electric had the ambition to be able to advertise using the phrase: "as alike as two telephones". But they found that, the harder they tried to achieve consistency and uniformity, the worse were the effects. The more they tried to *shrink* variation, the *larger* it got. When any kind of error, mistake or accident occurred, they went to work on it to try to correct it. It was a noble aim. There was only one little trouble: things got *worse*.
>
> Eventually the problem went to Walter Shewhart at the Bell Laboratories. Dr Shewhart worked on the problem. He became aware of two kinds of mistakes:
>
> 1. Treating a fault, complaint, mistake, accident as if it came from a special cause when in fact there was nothing special at all, i.e. it came from the system: from random variation due to common causes.
> 2. Treating any of the above as if it came from common causes when in fact it was due to a special cause.
>
> What difference does it make? *All the difference between failure and success.*
>
> Dr Shewhart decided that this was the root of Western Electric's problems. They were failing to understand the difference between common causes and special causes, and that mixing them up makes things worse. It is pretty important that we understand those two kinds of mistakes. Sure we don't like mistakes, complaints from customers, accidents; but, if we wade in at them without understanding, we only make things worse. This is easy to prove."

Unexpected as it might be, it seems that Dr Deming first used the terms "common cause" and "special cause" around 1947 in discussions on *prison riots* (see *Out of the Crisis* page 270[pages 314–315]). Did something *special* occur to spark off a riot? Or was the riot due to the procedures, the environment, the morale of both the prisoners and the prison staff, the way the staff treated the prisoners, etc, etc? That is, was the *common* state of affairs (which Deming would refer to as the *system*) in the prison such that riots would be *bound* to occur from time to time? (Think back to Day 1's Major Activity.) Or would it take something *special*?

Activity 3–a is also on Workbook page 31.

ACTIVITY 3–a

Suggest some of the possible causes of variation in the arrival-time of my school bus.

Unsurprisingly, I do not have any data available from all those decades ago from which we could construct control charts. So, with each of the causes you've suggested above, simply say whether you *think* the cause is likely to have been common or special, and why. (Be guided by the paragraph on prison riots that we have just seen on page 2.)

Being guided by the thoughts about prison riots, factors which result in somewhat random fluctuations from day to day would be interpreted as *common* causes, while relatively rare events which markedly affect the arrival time in a "one-off" manner would be interpreted as *special* causes. So e.g. if the bus is held up by a serious road accident or by freak weather conditions such as a flood or unusually heavy snowfall then we would regard these as special causes. On the other hand, normal fluctuations such as the number of people queuing at the bus-stops, the proportion of red traffic lights encountered along the journey, usual day-to-day variations in traffic-density, etc would be regarded as common causes.

Activity 3–b is also on Workbook page 32.

ACTIVITY 3–b

There is a saying that "Variety is the spice of life". This implies that *variety* is good. In these early pages of Day 3, we (and Dr Deming!) have been arguing that *variation* is bad.

Is there a conflict here? *(No, but clarify why not.)*

And how might *reducing variation* lead to *increasing variety*?

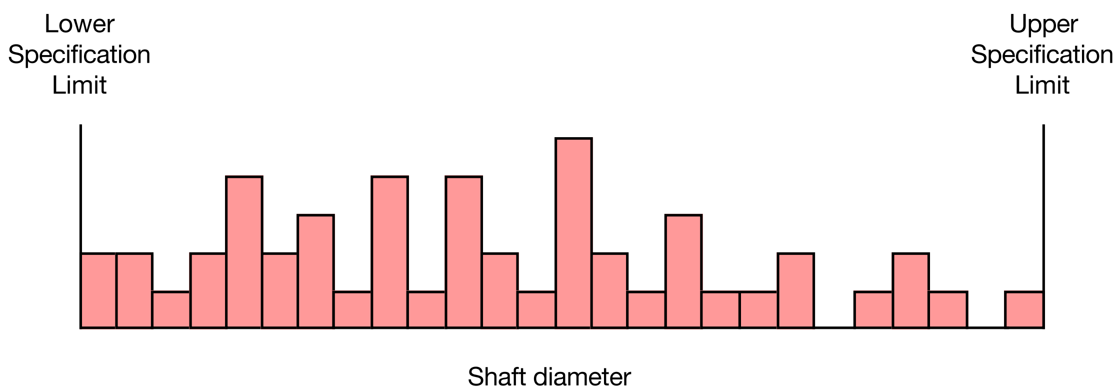
(For discussion see Appendix page 14.)

AT THE FORD MOTOR COMPANY

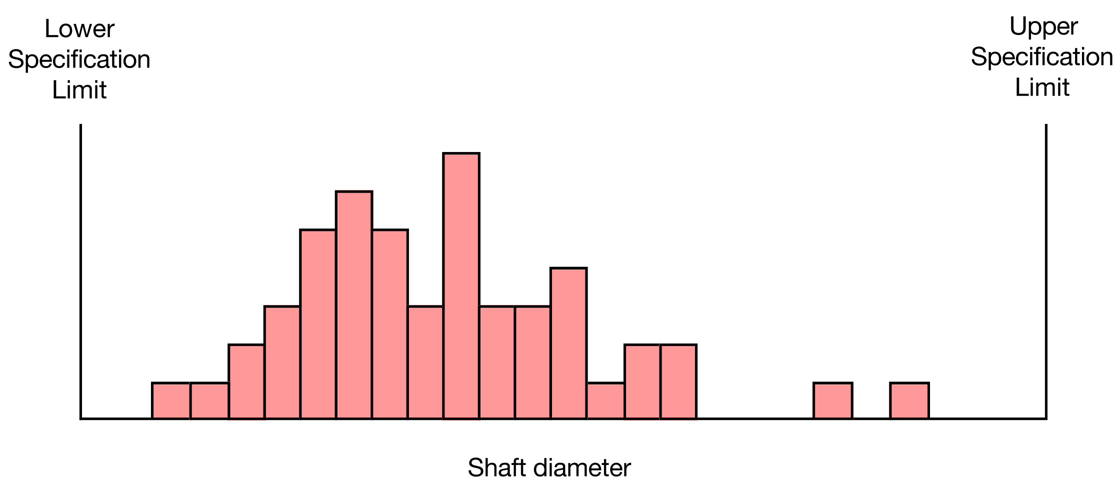
On page 2 Dr Deming claimed that, if we mix up the two types of variation, "we only make things worse. This is easy to prove." So let's prove it! Here is a well-known illustration, originating in the Ford Motor Company (see Bill Scherkenbach's *The Deming Route to Quality and Productivity*, around page 30 depending on the edition and reprint number[a]).

Input shafts for a transmission were turned in a machine equipped with an expensive automatic compensation device. If the diameter of a shaft was too large, the compensation device reduced the machine setting in order to try to make the next diameter closer to its nominal value. If the diameter was too small, the setting was increased, for the same purpose. Sounds sensible?

Here is a histogram of the diameters of 50 shafts consecutively manufactured by this process. (A "histogram" is a widely-used type of diagram for illustrating data. But in case you're not sure about what a histogram is, I'll guide you on how to draw one on page 7.)



A statistician recommended that the next 50 shafts should be made with the compensation device turned off. Here is the histogram of those 50 diameters.



Reduced variation! Better quality! Everything clearly within the specifications! *With that sensible and very expensive compensation device turned off!*

Activity 3–c is also on Workbook page 33.

ACTIVITY 3–c

To repeat:

> "Reduced variation! Better quality! Everything clearly within the specifications! *With that sensible and very expensive compensation device turned off!*"

How could this be?

(If you cannot answer this question now, be sure that you will be able to do so before you finish working through today's Major Activity!)

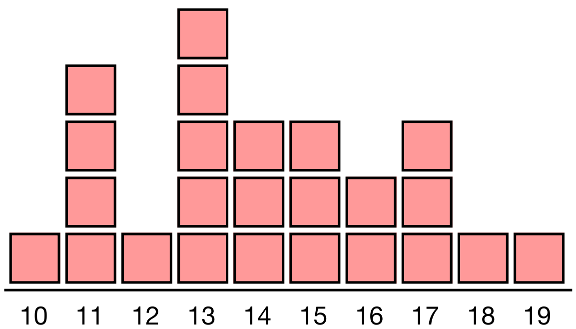
THE IMPORTANCE OF TIME

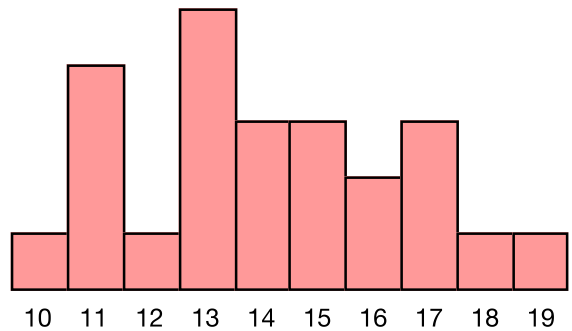
The histogram is often not a very informative method of representing data from *processes* (i.e. data having a natural order in *time*)—at least, not without accompanying it by a run chart or control chart. Why not?

This is a sequence of 24 values written down in the order generated by a process:

11 10 11 11 12 11 13 13 14 13 14 13 13 15 14 15 15 16 17 16 17 18 17 19

Here are two forms of histogram of those data. On the left, each item in the data is represented by a box stacked on the appropriate pile. Little gaps have been left there between the boxes so that you can see clearly where the boxes are. Such gaps are usually filled in, as is shown directly below and as in the Ford histograms you saw on page 5.



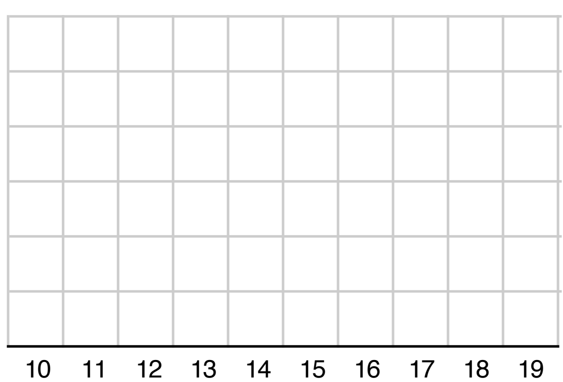


Activity 3–d is also on Workbook page 34.

ACTIVITY 3–d

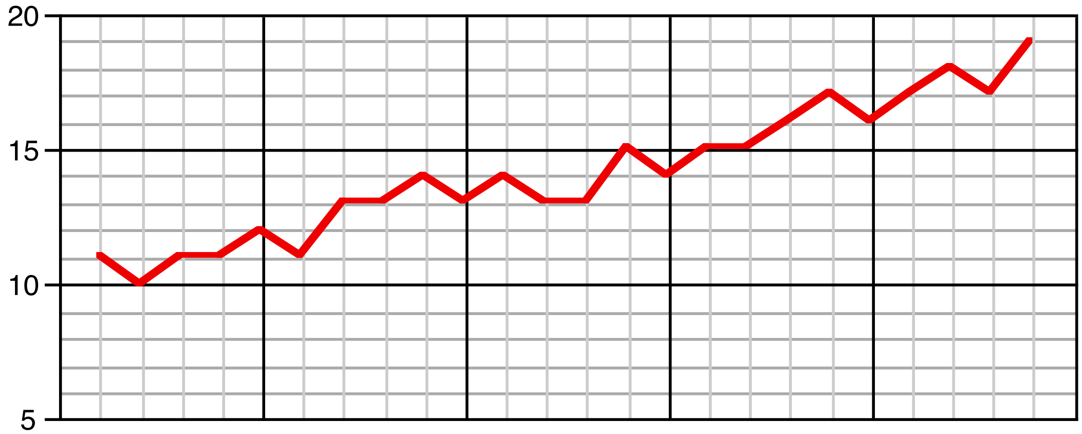
Here is another sequence of 24 numbers. Please sketch a histogram of these data. I suggest you use separate boxes as on the left above. What do you conclude?

18 19 17 17 16 17 16 15 14 15 15 13 14 13 14 13 13 13 11 11 12 11 10 11



A "similar histogram"? It turned out to be the very *same* histogram as that obtained with the first set of data! Yet the processes were surely very different from each other. That is to say: the *behaviours of the two processes over time* were very different from each other.

Just to be sure, let's draw run charts of the data. Here is a run chart for the first process.



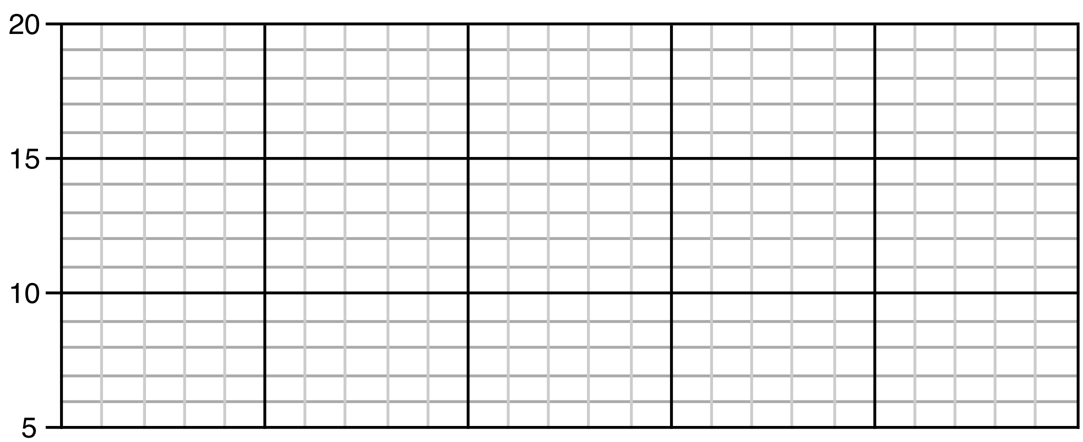
Activity 3–e (pages 8–9) is also on Workbook page 35.

ACTIVITY 3–e

Here again are the data from the second process (to save you from having to look back):

18 19 17 17 16 17 16 15 14 15 15 13 14 13 14 13 13 13 11 11 12 11 10 11

Please draw a run chart for this second process.



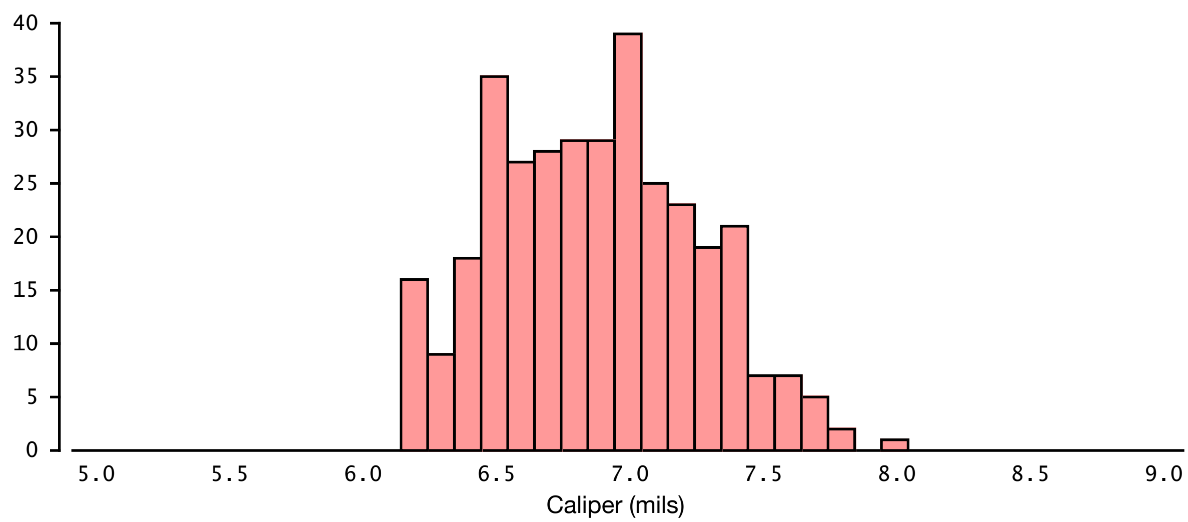
Compare and contrast the learning obtained from the two ways of pictorially representing data illustrated in Activities 3–d and 3–e.

Clearly, the first process was trending upward over time; the second was trending down. Very different behaviours. Yet we found exactly the same histogram in both cases. The two *collections* of numbers were exactly the same: they just occurred in a different order. But the *order* of the numbers is all-important for describing and understanding the behaviour of a process—it is foolish to ignore it.

Histograms totally ignore the order in which the numbers come out of a process.
Yet that order is very likely to hold the most important information of all about the process's behaviour.

However, that's not to imply we should dispense with the histogram altogether. The top priority is to learn whether or not a process is in statistical control. But the histogram can sometimes indicate important features which may be less clear on a control chart. For example, we have already seen how loudly the Ford histograms shout out the important message that the automatic compensation device was *increasing* the variation in the shaft diameters. That *could* also be seen from control charts, but I'd say the histograms demonstrate the contrast much more obviously and straightforwardly.

Dr Deming included a few histograms in *Out of the Crisis.* I particularly like his commentary on the following histogram which I have redrawn from page 229[267] of his book:



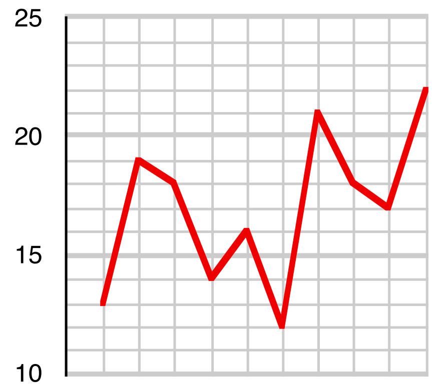
Observing the rather abrupt cut-off on the left hand side at 6.2 mils (millimetres), he wrote that the histogram …

> " … shows a distribution of measured values during production. The lower specification limit was 6.2 mils; no upper limit. No part recorded a failure. Note the peak at 6.2 mils. Were there any failures? No one will ever know.
> No one wishes to be the bearer of bad news."

Actually, if you look carefully at the first of the two Ford histograms on page 5, I think you will find that it contains not 50 but just 49 "boxes". Maybe there was one that fell just outside specifications and, well, … vanished.

MORE ON THE "SALES DATA"

Now, also as previously promised, let's return to the illustrative example with which we started Day 2. So, if you need to, take a quick look back at the first three pages of Day 2 to remind yourself about what happened there. We finished up with the run chart alongside of ten supposed monthly sales figures for a new product:



First, a short Pause for Thought:

Pause for Thought 3–f is also on Workbook page 36.

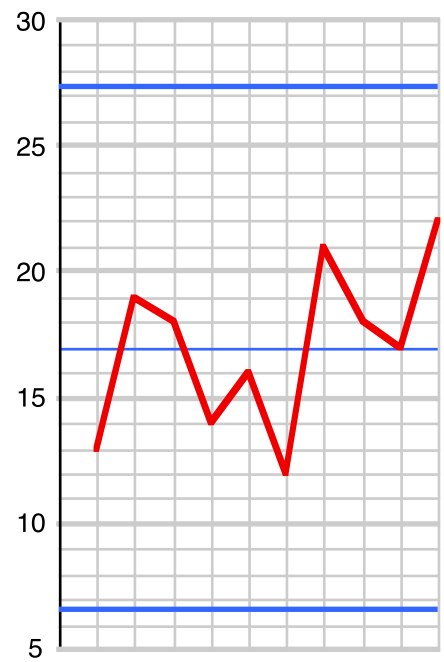
PAUSE FOR THOUGHT 3–f

Does the Ford example from earlier today (pages 5–6) suggest to you any concerns about the management's interpretation of yesterday's monthly sales data (Day 2 pages 1–3), the conclusions they drew, and the decisions they made?

This is not an exact analogy with the Ford example because, in that case, there was an "ideal" or "target" value, and the compensation device acted according to whether the value was above or below the target. Nevertheless, the management's behaviour here was somewhat analogous—generally acting one way if the figure went up and another way if it went down. The suspicion which the Ford example might therefore raise in our minds is that maybe such reactions could actually have *increased* the variation in the figures.

Furthermore, doesn't that type of management behaviour remind you of what you were doing yesterday when playing the part of the Foreman in the Red Beads Experiment? With the sales data, the management team interpreted every month's figure "in its own right". On some occasions it was bad news, requiring immediate and possibly expensive action. On other occasions it was good news and so they could breathe easily again. The Foreman moaned about high numbers of red beads and came up with "reasons" for them. He praised the low numbers of red beads and found reasons for them too. Did any of this make any sense from the All-Knowing's viewpoint? Remember that the All-Knowing genuinely understood *everything* there was to know about the process which was producing the data (an exceptional privilege!).

Let's gather together some thoughts. What might our basic learning about variation tell us? One crucial matter which is (a) obvious but (b) largely ignored by those who lack such basic learning is that all but the most trivial of processes exhibit *some* variation. Thus, as I said in my comment on yesterday's Pause for Thought 2–a (Day 2 page 4), the figure being recorded will indeed usually either go up or down! That being the case, greater attention needs to be paid not to *whether* the figure goes up or down but to *how far* the figure goes up or down. Now, the management team did this to an extent, but only to a relatively *small* extent. And, without the use of control charts, that is what usually happens. Even if people appreciate that processes do indeed have their own inherent variation (the *common-cause* variation) the truth is that, without using a control chart, they almost always underestimate how large it is (recall my first additional learning-point from the Red Beads Experiment on Appendix page 10). Thus they are *still* largely prone to the effect shown in the Ford example: increasing variation as opposed to doing anything useful. The only reliable way to assess the order of magnitude of the common-cause variation is to use a control chart.



So let's now carry on as we did through much of yesterday with the various sets of Red Beads data: let's now upgrade the run chart to a control chart by inserting control limits.

With that build-up, the control chart on the right should perhaps be less of a surprise than if I had shown it to you yesterday before you began learning from the Red Beads Experiment. Look where those control limits are in relation to the data!

Returning to the question at the bottom of the previous page (concerning the All-Knowing's viewpoint), did any of the management team's interpretations of the data, and did any of the Foreman's interpretations of the numbers of red beads in the Red Beads Experiment, make any sense to the *Un*-Knowing? Remember that, though knowing nothing about the process itself, the Un-Knowing did understand how to interpret a control chart!

All the data are happily contained within the control limits. The indication from the control chart is that, as in the Red Beads Experiment, this process is *stable*, in statistical control. According to the control chart, there is no evidence that the point-by-point interpretations of the data as carried out by the management team made any sense whatsoever.

But remember, as I emphasised in Pause for Thought 2–a, "these were not real sales figures, and this was not a real management team". Why did I use artificial data rather than real sales data? That will now become clear.

> As a minor point of detail, notice the thin blue line in the middle of the chart. This is the "Central Line" representing the average of the data-values from which those control limits have been computed. More often than not, people do include the Central Line on their control charts. Dr Deming did not bother with the Central Line when control-charting Red Beads data, and neither did I. It is nowhere near as important as the control limits. Nevertheless, following common practice, we shall usually include it from now on.

Here I am able to become the "All-Knowing"—because I know where these data originated. In fact, if you happen to have a copy of the second edition of my *Elementary Statistics Tables* (abbreviated on Day 1 by *EST*) then you might be able to find them for yourself! For those data are in fact the first ten of the 50 data from one of three case studies included there (see *EST* page 48, Figure 1). Yes, those data do come from a process, though not a sales process. They come from a process of … *throwing dice*! Those ten data are the total spots which showed when I threw five ordinary dice.

> These really *were* “honest” data—I didn’t keep on throwing the dice until I obtained a sequence which would make a good story! Indeed, it didn’t even occur to me to use these data for such an illustration as this until some considerable time after the new edition of *EST* was published.

Apologies if all this seems to have been something of an elaborate hoax—although I have given you plenty of clues that all might not be as it seemed! There was a very important point for me to re-emphasise following yesterday’s work on the Red Beads Experiment, and this seemed a good way of doing it. It’s simply this:

If your data are typical of what you would get if you were just throwing dice (or using some other similarly “random” mechanism), what justification is there in trying to interpret individual figures from that process?

Try to argue with this if you will—but, if these data are all that you have, the answer is: *none*! And this is precisely the kind of sensible conclusion that the control chart encourages you to make, and then helps you to make it.

Incidentally, you may have noticed that every control chart you have seen so far has indicated that the relevant process is stable. Maybe you have begun to think that *all* control charts show statistical control! This will be immediately remedied at the start of the “Six Processes” section on page 19.

NB We have now reached the stage where I strongly advise Stats-level 0 students and, even more so, Stats-level 00 students *not* to work hard at the following pages. Even higher Stats-level students will be hard-pressed to take in everything on these pages at a first reading. So, really, I definitely recommend *everybody* to spend *no more* time here than the clock icons advise. Just get a *flavour* of what’s here for now, and then return to study it more carefully at some time in the future when you start collecting and analysing your own real-life data. The material on pages 13–34 is effectively *extra-curricular*—it is *not* used anywhere within the course after today. But I believe it *will* prove valuable to you when the time comes for you to start constructing and interpreting your own control charts.

HOW DO WE COMPUTE THOSE CONTROL LIMITS—AND WHY?

Let me remind you of something I told you when first referring to the Red Beads Experiment (Day 1 page 7):

> "[Dr Deming] would usually draw up a [control] chart once the results were obtained in his famous Red Beads Experiment … But how? He would just write down a simple formula, insert some numbers obtained from the experiment, and do the arithmetic. But there was nothing about where the simple formula came from nor the fact that, with the large majority of processes, that same formula wouldn't even be appropriate!"

Further, during Technical Aid 1 on Day 2 page 20, my introduction to that method for computing control limits was as follows (slightly abbreviated):

> "One of the earliest applications of Shewhart's invention of the control chart was for batch inspection of mass production processes. In such inspection, samples (batches) of items from the process's output are regularly drawn and inspected, and the number of defective items recorded."

Interpreting the red beads as "defective items", we immediately see that the method for computing the limits described there is indeed appropriate for Red Beads data. But we must be clear that the theory underlying that method depends upon

> [A] Shewhart's guidance on where to place the control limits

and

> [B] some particular characteristics of the behaviour of data from batch inspection processes

—*both* [A] *and* [B]. Now, [A] is very generally applicable to different kinds of processes that we are concerned with in practice but, of course, [B] is not. In particular, neither real sales data nor data generated by throwing dice match the "batch inspection" model in any way. Without [B], the "batch inspection" method of computing control limits makes no sense, has no relevance. Instead we need a method which uses [A] but does not have [B]—or maybe anything like [B]—available. So that is surely true for the control limits shown on page 11: the data there have *nothing* to do with "batch inspection". Simply stated, we need a method for computing control limits which *only* uses Shewhart's guidance plus some data from the process that we want to study—nothing else: no other information, no other assumptions. But, in principle, this presents a serious logical dilemma. Why?

As we know, the control limits need to indicate the range over which the data will vary when the process is in statistical control: so that, if and when data go outside those limits, we have evidence that the process may well be out of statistical control. That is, of course, quite feasible *if* the process is in statistical control when we collect the data from which the limits are computed. But that begs the question. Suppose the process is *out of* statistical control when we collect those data. Surely that same method is then likely to produce control limits which instead indicate the range of the data when the process was *out of* statistical control! To put it mildly, that's not much use! So surely we need to check that the process is in statistical control when we collect our data. But in order to do that we need a control chart. Ah, but we don't have one yet—that's what we are trying to produce. Wouldn't *you* call that a "serious logical dilemma"?

So ideally we need a method for producing control limits which indicate the extent of common-cause variation, whether *or not* the process is in statistical control during the time that the data are being collected from which we shall compute the limits. That's a tall order, and there is no perfect solution. In particular, if using the type of methods for measuring variation which are familiar in traditional Statistics (particularly the "standard deviation"), the effect of special causes is usually to considerably *widen the gap* between the control limits. That will, of course, destroy the ability of the control chart to detect special causes, effectively making the chart useless.

But there is a better way. There is an ingenious yet simple approach to substantially reducing the "contamination" effect of special causes on the control limits. It harks back to a previous section: "The Importance of Time" on pages 7–9. The standard deviation and similar traditional measures of variation pay absolutely no attention to the *order* in which the data are generated (which, recall, is incidentally the same disadvantage as that suffered by histograms). In fact, if you used them to produce control limits for the data presented in the "Importance of Time" section, the control limits would actually be so far apart that all the data there would be contained well inside them—despite the fact that those data hardly look as if they came from a stable process such as in the Red Beads Experiment! (For if they did then the Foreman would *really* have had something to get excited about!)

So the approach that we use instead measures variation is a completely different way—a way which *is* wholly dependent on the *order* in which the data are generated. So let's be clear about the fundamental difference between the two approaches:

- The traditional measures of variation typically focus simply on how far away the values in the data are from the *average* of the data-values

whereas

- The method we are now introducing is based on the point-to-point *changes* in the data, i.e. how far away each data-value is from the data-value that preceded it in time-order.

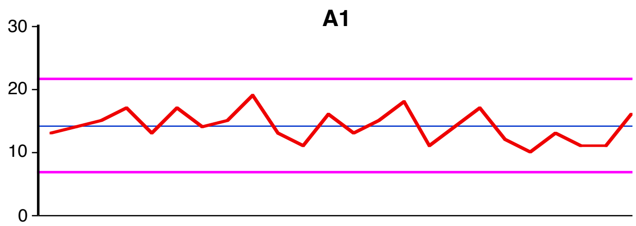
These point-to-point changes, i.e. differences between consecutive values in the data, are known as "moving ranges". Obviously, if many of these moving ranges are large then high variation is indicated; whereas if the moving ranges are mostly small then low variation is indicated. So this is certainly a *sensible* alternative method of measuring variation. But that is a relatively minor point.

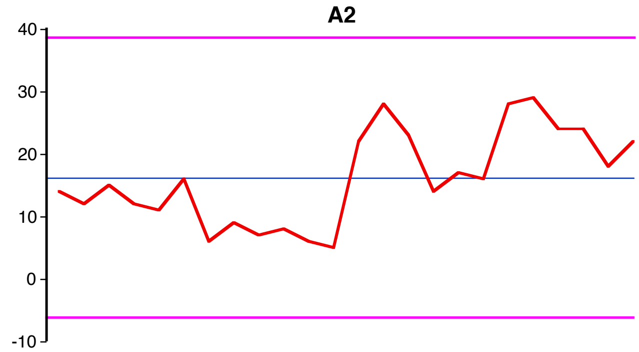
The major point is that, in most cases, using moving ranges *substantially reduces* the "contamination" effect of special causes on the control limits compared with using the standard deviation or anything similar to it. Suppose, for example, that a special cause simply raises or lowers the process *average* and that, unfortunately, the numbers used for calculating the control limits include data from both before and after that change; the Central Line and control limits are then likely to move quite a lot from where they would have been otherwise. The vital advantage of using moving ranges to compute the control limits is that then it's often the case that the special cause alters *the distance between* the limits by only a small amount—so that the chart *retains* its sensitivity to detect special causes. That is *not* true if using traditional measures.

How does the moving-range method achieve this? Let's consider the case where a fault occurs which *raises* the process average from what it was before. That fault thus also similarly raises all of the subsequent data—but *not* their variation. That special cause will in fact only affect a *single* moving range: the one starting with the data-point immediately before the special cause occurs. All the *other* moving ranges are unaffected by the change in process average since they all compare two data-values which were *both* recorded when the process had the original average or were *both* recorded when the process had the raised average. Thus, other than that one exception, all the moving ranges *still* continue to reflect the size of the process's *common*-cause variation. So the computed control limits will be *slightly* contaminated by the special-cause effect (because of the one exceptional moving range) but usually not to the extent of seriously harming the control limits' ability to do their job: that of indicating the presence of special causes.

This kind of feature is so important that we'd better study some charts. Firstly, take a look at the two control charts, labelled A1 and A2, at the top of page 19. *(Recall that on Appendix page 1 I indicated you would find it convenient to print a separate copy of that page.)* I think you will quickly realise that Chart A1 shows a process which is in control. But Chart A2 looks very different. Amongst other features, four of its 24 points are very close to, or outside, the control limits. Thus there is little doubt that the data shown there come from a process which is *out of statistical control*.

The control limits (coloured blue) in those charts on page 19 were computed by the *moving-range* method (in both cases using all of the 24 data-points included on the chart in the calculations). However, what would happen if we were instead to try computing the control limits with the conventional statistician's *standard deviation* (again using all the 24 data-points on each chart)? The control limits (coloured purple) on the charts that follow are computed using standard deviations.





Now, those purple limits on this Chart A1 (computed using standard deviations) are virtually the same as the blue ones (computed using moving ranges) on page 19. *But* look what happens to Chart A2 where the standard-deviation-type (purple) control limits have been computed from the data there *which we have already just recognised as coming from an out-of-control process*. They are approximately twice as far apart as the ones on page 19! *These* limits are *totally useless* as a criterion for judging whether the process is or is not in statistical control. *Why* does this happen? It's because, just like the histogram, the standard deviation totally ignores the *order* in which the numbers occur in the data—refer again to the remark about "the most important information of all" near the top of page 9.

Now, as I said, using moving ranges is not a *perfect* solution for obtaining control limits that purely reflect common-cause variation even if computed when the process is out of statistical control. A perfect solution doesn't exist. But using moving ranges works pretty well in mitigating the contamination effects of many kinds of special causes. There are a few exceptions but, now that you know the general idea about how the control limits are produced, i.e. using moving ranges, those exceptions soon become relatively easy to recognise. That is why I have included this present discussion here in the main text rather than just in the Technical Aids—this is important knowledge for *all* users of control charts, including those on Stats-level 0. Two important exceptions that one needs to be able to recognise are illustrated with data generated in the Funnel Experiment, and so we shall see those this afternoon. But the general success of the moving-range method will be amply illustrated this morning in the section which begins on page 19.

Those on Stats-level 0 can now move to that section almost immediately since the remainder of this current section consists of some Technical Aids and an Activity in which to use those Technical Aids. However, since most of this afternoon will be spent on the substantial Major Activity of carrying out a version of the Funnel Experiment, there will then be no further formal Activities or Pauses for Thought during the rest of this morning. The important "activities" related to what follows this morning will instead be those which take place *in the future* when you start to interpret your *own* control charts using your own data from processes that are of interest and relevance to you!

Let me re-emphasise that what follows, all the way up to page 34, is effectively "extra-curricular" as far as this course is concerned. Neither this afternoon's work nor anything during the rest of this course will depend on the content of those pages. Further, the substantial Optional Extras material that has been previously mentioned is specifically for those who want to gain both deeper and broader knowledge about control charts than is included in the main course, particularly including more technical details. So that will not be suitable for everybody—which is, of course, why it is indeed "Optional Extra"! By contrast, pages 19–34 here are focused on helping you to *interpret* control charts: they are not concerned with more technical matters nor with the actual *construction* of control charts. So, although they are optional *as far as the rest of this course is concerned*, I believe you will find them extremely helpful if and when you become actively involved in control-charting.

I hope therefore that you will find pages 19–34 interesting to browse through now (so that you get some idea of what's there) but, more importantly, that they will then become ones to which you will return from time to time, particularly when you use control charts in your own work and elsewhere. There is much reading and thinking involved during these coming pages. ***Please do not expect to take it all in straight-away today—there's a lot if it!*** Keep a careful watch on my timing-guidance. This is the one and only occasion when I encourage you to just "skim-read" a substantial section of the main course material.

On the other hand, if you find it too difficult to deal with everything in the time that I've allotted to these pages then there's no harm done. When the time comes for you to start working with your own control charts then you can return to study at your own pace the guidance given in these coming pages. However, please note that, if you do decide to skip some (or all!) of these pages now, be sure not to miss out pages 35–37 since they cover essential preparation for this afternoon's Major Activity and thus are *not* optional!

So, if you are on Stats-level 0 (or 00), please omit the following Technical Aids and move on to page 19.

Technical Aid 5

With Red Beads and similar data, the σ in Shewhart's guidance for control limits (referred to on Appendix page 4) was computed by a formula merely involving the average: $\bar{X}$. That was possible because, with such data, there is a natural link between the average and the way the data vary. (If the average is particularly small *or* large then the variation is relatively small, whereas if the average is more central then the variation is relatively large.) This is *not* the case with most other types of data, and so then a more direct representation of the variation is needed. As has now been introduced, the recommended method uses *moving ranges*: the distances (positive or zero, *not* negative) between consecutive values in the data.

Technical Aid 6

The artificial sales data on pages 10–12 were, in time order: 13 19 18 14 16 12 21 18 17 22 .

(a) As with the Red Beads data, $\bar{X}$ represents the average of all the data that are being used to compute the control limits. We'll use all ten of them. So here we have

$$\bar{X} = (13 + 19 + 18 + 14 + 16 + 12 + 21 + 18 + 17 + 22) \div 10 = 170 \div 10 = 17.0\,.$$

This is where the Central Line of the control chart on page 11 was placed.

(b) $\overline{MR}$ represents the *average moving range* in the data. In what follows, the moving ranges are shown in italics. Note that the moving ranges are the *sizes* of the point-to-point changes: there are no minus signs involved. Also note that, with 10 data-points, there are of course just 9 moving ranges.

13 19 18 14 16 12 21 18 17 22
6 1 4 2 4 9 3 1 5

So $\overline{MR} = (6 + 1 + 4 + 2 + 4 + 9 + 3 + 1 + 5) \div 9 = 35 \div 9 = 3.889\,.$

(c) The control limits are placed at a distance of $2.66 \times \overline{MR}$ either side of the Central Line. Here we have $2.66 \times \overline{MR} = 2.66 \times 3.889 = 10.34$, and so the control limits were placed at $17.0 - 10.34 = 6.7$ and at $17.0 + 10.34 = 27.3$.

Technical Aid 7

Why the 2.66? Sorry: as with the formula in the Red Beads case, this is more material for the Optional Extras. But trust me: it *is* derived using the guidance about control limits provided by Dr Shewhart.

However, although moving ranges are quick and easy to compute, they can become quite tedious if being computed (especially by hand) for a *lot* of data. For example, it would be bad enough if we had to compute control limits using all the 24 data in the illustration below. Much worse still would be control limits for the data which you will generate in this afternoon's Major Activity: four lots of 40 numbers!

Technical Aid 8

If we want to turn run charts containing a lot of data into control charts, do we have to use all those data to compute the control limits?

Fortunately, no. Sometimes we do use all of the available data to compute control limits *retrospectively*, i.e. when studying past behaviour of a process (as in the following section and also in the Springboard article previously referenced). But otherwise it is more usual and useful to develop a control chart "live", i.e. plotting the points one at a time as and when the data are received. The normal practice is to compute control limits from, say, the first 10 to 15 data-values—or less if the data are received weekly or monthly, or if you're in a hurry to get started! The length of data used for calculating the control limits is sometimes called the "baseline". Of greater significance than just the convenience of using fewer data is that, obviously, a relatively short baseline can result in special causes being detected earlier than otherwise. There is much more discussion about this important matter of short or longer baselines in Part F of the Optional Extras section. There is also some discussion on pages 82–84 in *ST*, the revised edition of my *Statistics Tables*.

I'll illustrate Technical Aid 8 using the first set of 24 data from page 7:

11 10 11 11 12 11 13 13 14 13 14 13 13 15 14 15 15 16 17 16 17 18 17 19 .

Let's compute the control limits using just the first half of these data, i.e. the first 12 values. Following Technical Aid 6 we firstly compute $\bar{X}$, the average of all the data which are being used to produce the control limits:

11 10 11 11 12 11 13 13 14 13 14 13 .

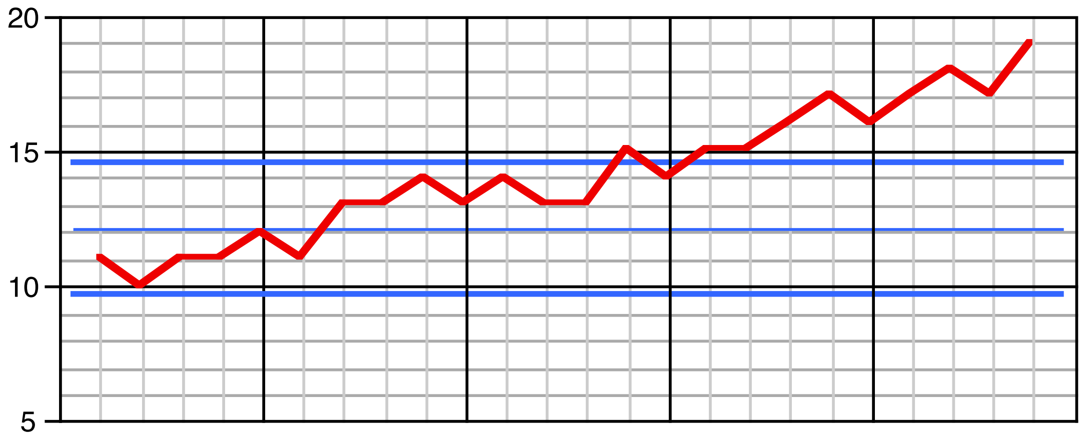
This is

$$(11 + 10 + 11 + 11 + 12 + 11 + 13 + 13 + 14 + 13 + 14 + 13) \div 12 = 146 \div 12 = 12.167.$$

Then it's the turn of $\overline{MR}$, the average moving range in those 12 data. As before, the moving ranges are printed in italics:

11 10 11 11 12 11 13 13 14 13 14 13
1 1 0 1 1 2 0 1 1 1 1

So $\overline{MR}$ = (1 + 1 + 0 + 1 + 1 + 2 + 0 + 1 + 1 + 1 + 1) ÷ 11 = 10 ÷ 11 = 0.909, and 2.66 × $\overline{MR}$ = 2.66 × 0.909 = 2.418. This puts the control limits at 12.167 – 2.418 = 9.749 and 12.167 + 2.418 = 14.585. Showing the control limits in blue as previously, the control chart is then:



This was clearly a case where, in practice, there was little to be gained by producing the control chart: the run chart had already told the story of the rising trend. That's fine: *if the run chart tells you all you need to know then don't bother to upgrade it to a control chart*. The control chart becomes really valuable when it is *unclear* as to whether or not the run chart is indicating there are some special causes—which is the more usual situation.

Activity 3–g is also on Workbook page 37.

ACTIVITY 3–g

Just for practice, compute the control limits by the method just demonstrated (again using just the first 12 values) on the data whose run chart you drew in Activity 3–e on page 8. Here are those data:

18 19 17 17 16 17 16 15 14 15 15 13 14 13 14 13 13 13 11 11 12 11 10 11

(Hint: I deliberately chose these numbers to provide easy arithmetic for you—in particular, you should find that both $\overline{X}$ and $\overline{MR}$ computed from the first 12 values turn out to be whole numbers.)

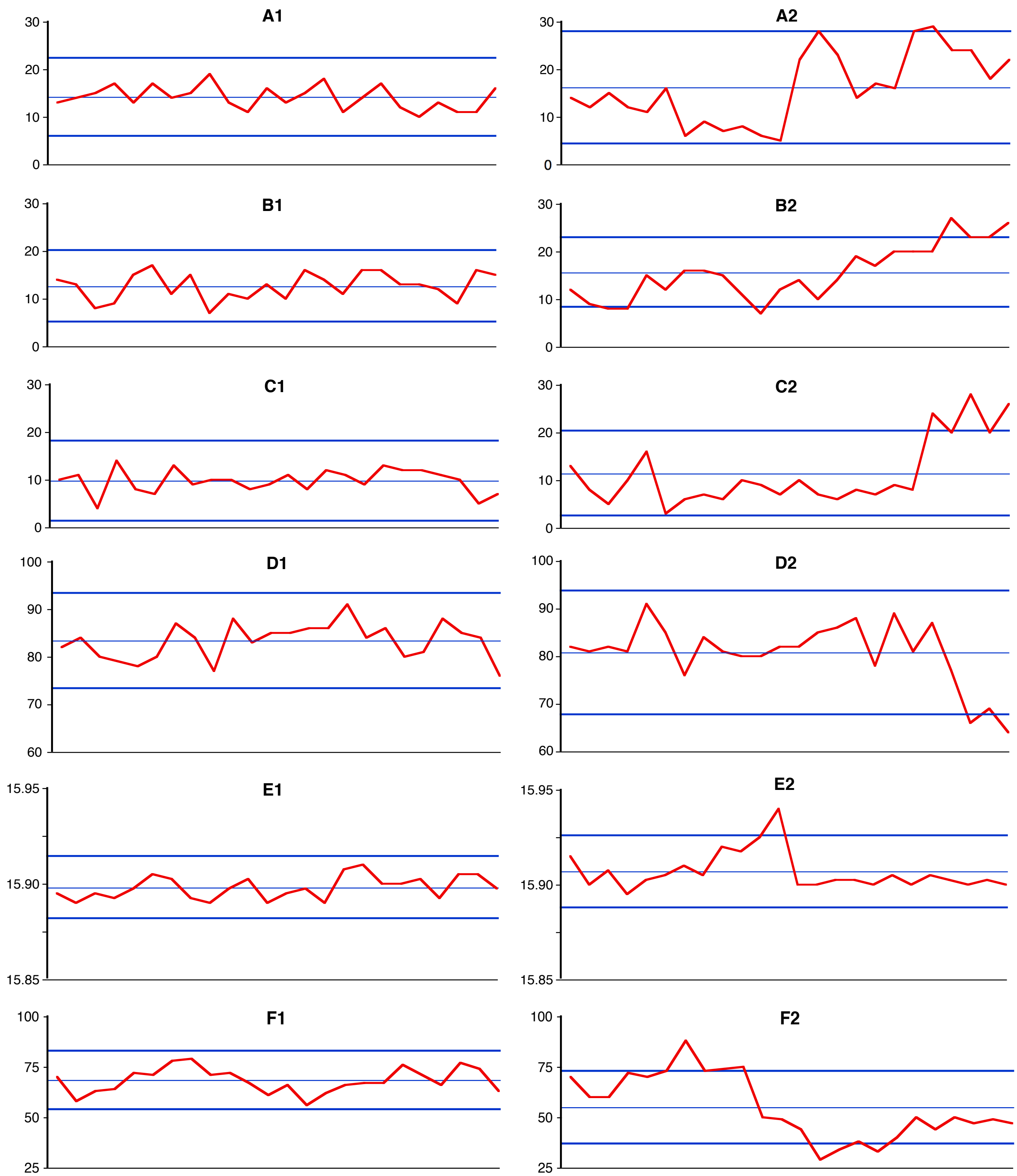
(If you need to check your arithmetic then see Appendix page 15.)

Then insert the control limits on your run chart on page 8.

If you had needed the control limits to help you interpret the data, what would they have told you?

SIX PROCESSES

Take an initial look at these twelve control charts:



To simplify comparisons, each of the 12 charts was constructed from the same number (24) of data-points and in each case the control limits have been computed from all of those 24 values. You will probably guess why I have chosen those particular details. Yes, one of the processes charted here is the one with which you are currently most familiar: a Red Beads process; and those were precisely the details relevant to all of yesterday's illustrations. (This choice of baseline length does not contradict the advice given in Technical Aid 8 since these analyses are all retrospective rather than "live".)

Six of these charts indicate that the relevant processes were in statistical control. The other six charts indicate the relevant processes were out of statistical control. I think you will easily be able to see which are which! Also, both from the title of this section and the way I have numbered the charts, you will probably guess that the 12 charts are formed in six pairs, the charts for each pair both coming from the *same* process. On the left hand side (Charts A1–F1) the charts indicated that the processes were in statistical control. On the right hand side (Charts A2–F2) the data used were recorded when all the same processes had gone out of statistical control.

I collected together such a set of twelve charts a *long* while ago and always found them to be very useful for beginning to get my delegates and students used to interpreting control charts. The charts cover a very broad range of types of process; I'll give you some details a little later on. But before getting into those details, I'll make some general observations.

One aspect almost immediately noticed by the delegates in my seminars is how control charts of stable processes (in statistical control) all look broadly similar to each other, irrespective of what type of processes they are. Naturally, Charts A1–F1 differ in detail but they all give the same kind of general impression. That should be no great surprise if you recall the description of data from stable processes on page 12: such data "are typical of what you would get if you were just throwing dice (or using some other similarly 'random' mechanism)". So, in reality, nothing of interest happens in such processes. In contrast, Charts A2–F2 differ in much more than mere detail, reflecting the fact that special causes are affecting the way that these processes are now behaving and can do so in all sorts of different ways.

Related to this is that the delegates often pointed out that Charts A1–F1 looked pretty "boring" whereas Charts A2–F2 looked relatively "interesting". Again, those impressions are not surprising: processes that are in statistical control produce data whose general behaviour stays the same. So no wonder they might be regarded as looking "boring": they are all "much of a muchness". Similarly, why do Charts A2–F2 look "interesting" rather than "boring"? Because, in each case, something has (or some things have) happened in them—indicated in particular by one or more points lying outside the control limits. These points warn us of real changes, caused by something different from the routine factors which were previously the only influences on the process. Something has changed which results in changed behaviour—usually, in practice, *worse* behaviour of the process.

So notice that (at least as far as control charts are concerned!) "boring" is *nice*! Charts A1–F1 tell us that the processes are in statistical control, stable, *predictable*. It is *nice* to have some idea of what our processes are likely to produce in the future—the near future, at least, i.e. until something occurs (intentionally or unintentionally) to change matters. The prediction is that, as long as the process stays stable, future data will continue to behave in the same manner as they are currently behaving. To put it mildly, that is *useful* to know!

On the other hand, of course, "interesting" is usually *not* nice! Instead it means there are some problems that need to be solved. But at least the control chart may provide some clues which may help us to solve those problems.

To see how, let me now tell you what the six processes were and what happened to them.

They divide equally into two sets. The first three are ones where I can again become the “All-Knowing”: one is a Red Beads Experiment and the other two are illustrations with dice and tossing coins. The other three are definitely “serious” processes where neither I nor anybody else could claim to be “All-Knowing”.

It is Process C which is the Red Beads process: the data are, as usual, the numbers of red beads finishing up in the paddle. The data for Process A are the total “spots” showing when four dice are thrown. The data for Process B are the numbers of Heads when 25 coins are tossed. So, similarly to the Red Beads Experiment, the dice in Process A and the coins in Process B are respectively thrown or tossed 24 times.

Regarding the “serious” processes, where nobody is “All-Knowing”, the charts for Process D show my pulse-rate taken at breakfast-time over a period of 24 consecutive days. The charts for Process E show average measurements in 24 small samples of manufactured cigarette-lighter sockets in a Japanese case study. There will be some detailed description of those measurements on page 32. Finally, the charts for Process F show the monthly United States trade deficits in billions of dollars over a two-year period. As I said earlier, these charts “cover a very broad range of types of process”!

In the case of Processes A–C, the above descriptions apply strictly only to Charts A1–C1. Clearly, if they also applied to Charts A2–C2 then I would have had some difficulty in finding any data which would indicate the processes were out of statistical control! So in those cases I deliberately *introduced* some special causes. I will also tell you what I know about the circumstances underlying the data in Charts D2–F2, over which of course I had no such direct influence!

For Chart A2, four dice were used in the first six throws as was the case throughout Chart A1. But only two dice were used in the next six throws and then six dice for the rest of the time. Similarly, Chart B2 began by tossing 25 coins as was the case throughout Chart B1, but I increased the number of coins by two each time from the 15th point onwards, finishing up with 45 coins by the end of the chart.

However, as you would expect, it was pretty difficult to figure out how to upset the Red Beads data! All I could think of was (for the final six points) to add up the two junior inspectors’ counts rather than to plot their common value (sorry—rather feeble, I know!). So, in that case, it wasn’t the process itself which went out of control: it was the process of *recording the data* that was in trouble. But that’s also important. Deming often pointed out that the *measurement* process needs to be in statistical control as well.

Moving on to the “serious” processes, you may have noticed that my pulse-rates on Chart D1 were rather unhealthily high. Chart D2 shows my pulse-rates for a later period of 24 days with the final four days showing the effects of a newly-prescribed beta-blocker. In the Japanese case study, a fault developed during the period covered by Chart E2, a fault which was soon more than effectively rectified.

Process F was the only one whose details I have changed from my initial version of the six processes. Despite the more than 30 years that have passed since then, it seems to me that the first five processes have stood the test of time pretty well. The original version of Process F consisted, as now, of monthly data on US trade deficits. However the years concerned then were 1988–89 for Chart F1 and 1990–91 for Chart F2. That earlier version of Chart F1 showed stability but Chart F2 showed some temporary instability, presumably because of the relatively minor recession which officially lasted from July 1990 to March 1991. However, because of the antiquity of those data and the considerably more serious international financial crisis of 2008, I thought it might be interesting to instead try the US trade deficit figures for 2006–07 for Chart F1 and 2008–09 for Chart F2. I was not disappointed!

The Springboard article also includes some discussion on these six processes except that there the *earlier* versions of Charts F1 and F2 are shown. On page 33 I’ll compare what happens in the two versions.

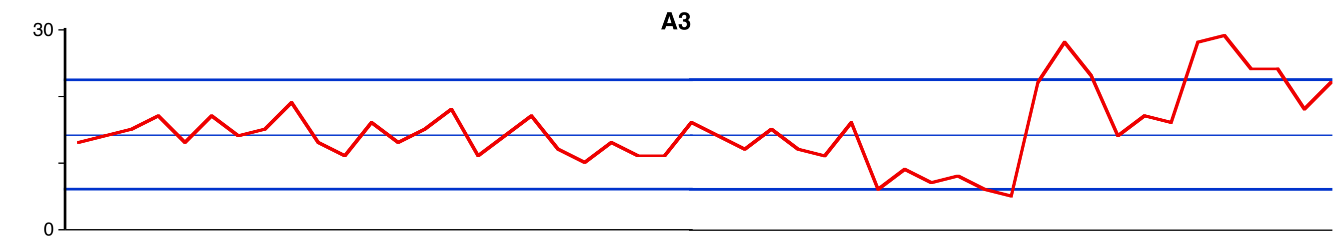
There is yet more to be read from these charts. However, I shall discuss just one further important issue at this stage and then return to these processes later on.

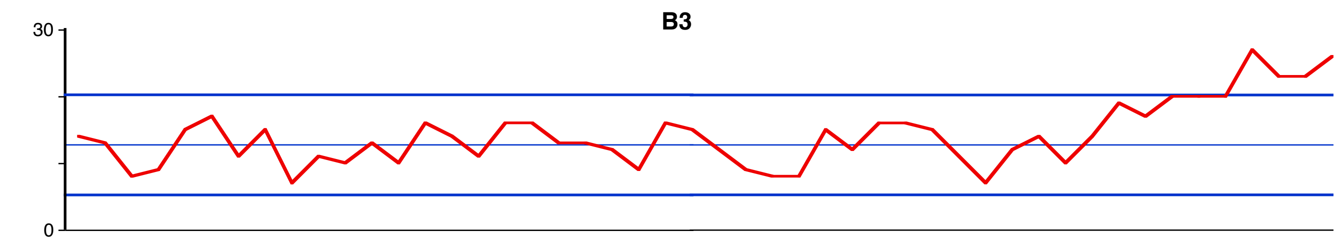
In every case, the in-control chart on the left-hand side of page 19 involved data which occurred *prior to* the data for the out-of-control chart on the right-hand side. Recall that one of the important interpretations of processes being in statistical control is that they are *predictable*. To repeat an important sentence from page 20, "The prediction is that, as long as the process stays stable, future data will continue to behave in the same manner as they are currently behaving." This surely implies that, if we have a control chart which indicates the process is currently in statistical control, it is sensible to extend the current control limits into the future.

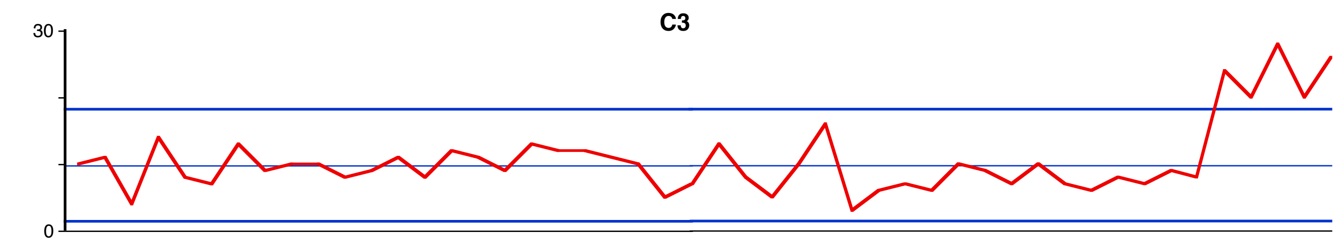
There are two advantages of doing this. First, it saves you time and effort by not recomputing the control limits when there is no need to do so. Second, if a future recomputation of control limits is being carried out while the process *is becoming* unstable, you'll obviously be spoiling the control chart's chance of warning you about that particular problem. And that's in addition to the likelihood of widening the control limits because of contamination from that special cause. So the clear "no-brainer" message is to leave the control limits alone if you have no good reason for changing them!

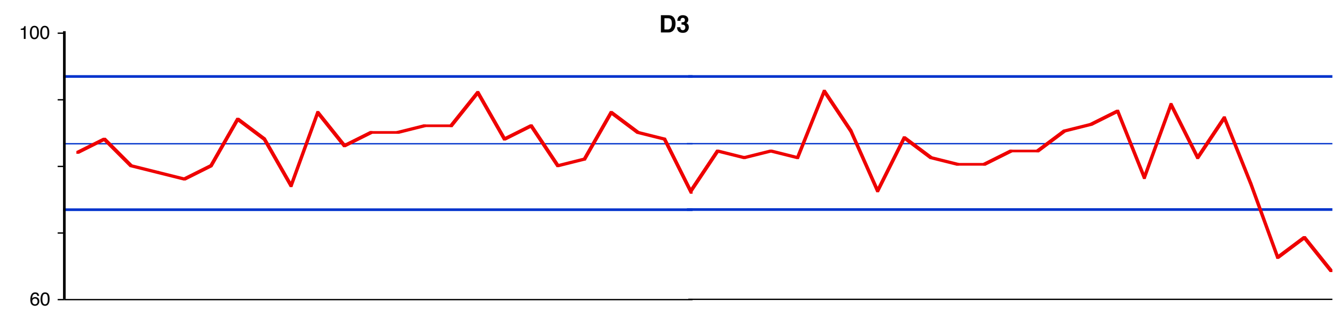
To illustrate these matters, on page 23 I have constructed Charts A3–F3. They contain respectively all 48 data from Charts A1–F1 and Charts A2–F2 but show the control limits from Charts A1–F1 throughout, i.e. extended into the future from when the first set of charts ended. If you compare Charts A3–F3 with Charts A2–F2 on page 19 then you will see that the signals of instability are generally stronger, i.e. the relevant points are mostly further outside the control limits than they were when the control limits were computed using the new data.

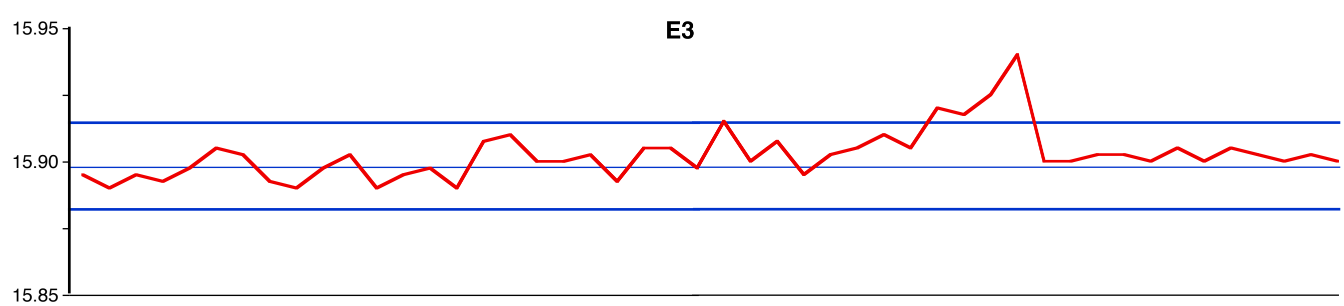
Thus note that we have obtained *more useful* results by doing *less work* (in this case, by not recomputing the control limits unnecessarily). This is a message that will recur during the Funnel Experiment this afternoon. It reminds me of Dr Deming ruing a sad fact which he frequently observed, namely: "people working hard, doing the wrong thing".

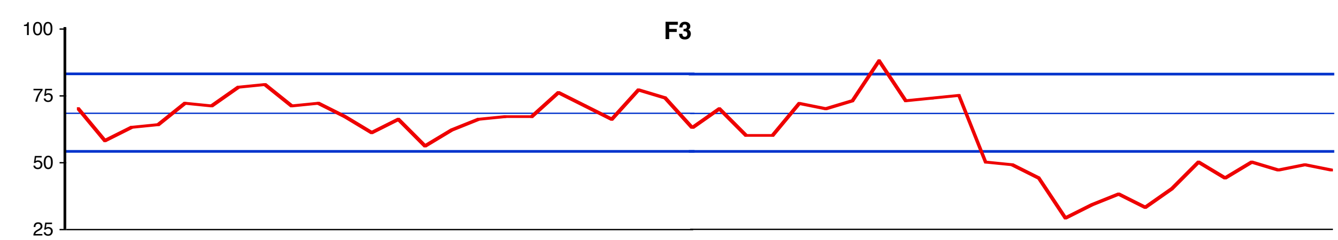

A3
30
0


B3
30
0


C3
30
0


D3
100
60


E3
15.95
15.90
15.85

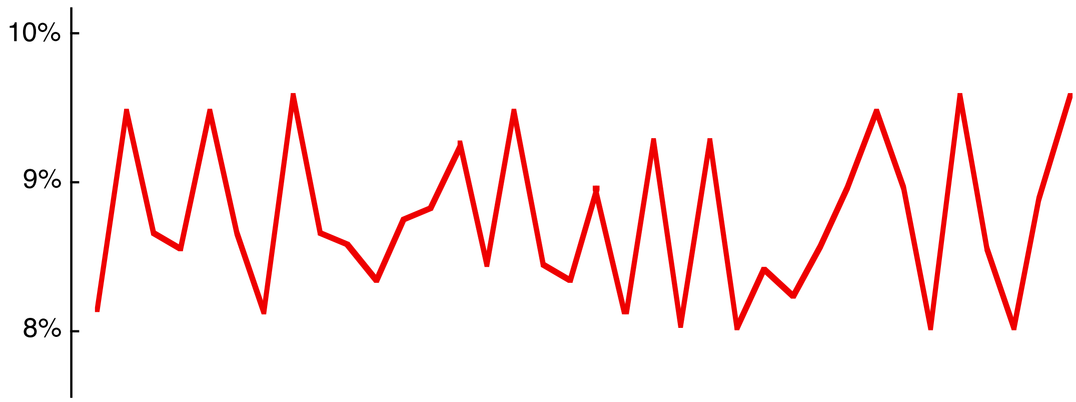

F3
100
75
50
25

A FAVOURITE EXAMPLE

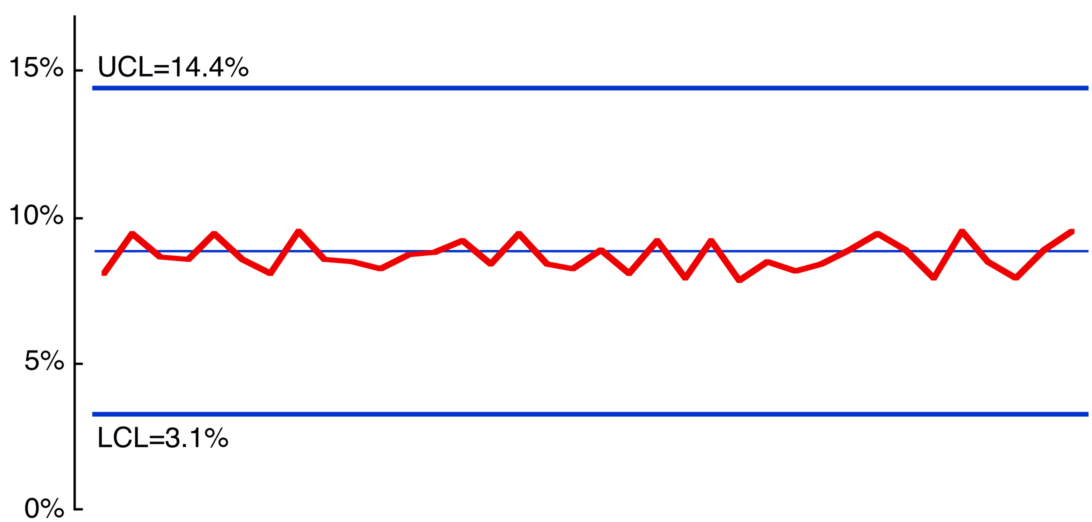
Through our study of the "Six Processes" we have already ventured well beyond the mere basics of interpreting control charts. So let's continue the good work a little while longer.

The example described below is one which Dr Deming frequently showed in his four-day seminars and is covered on *Out of the Crisis* pages 227–228[264–266]. Instead of simply reproducing what is shown in the book, I have redrawn a couple of charts based on the information in his account.

The following run chart is based on some inspection data. Quoting from Dr Deming's account, it "shows the daily record for two months of the proportion defective found on final audit of a product ready to ship out." It is quite usual, as here, to chart the *proportion* or *percentage* of defectives rather than the actual *number* of defectives—the run chart is exactly the same in each of the three cases except for what is written on the vertical axis. Of course, if upgrading the run chart to a control chart then the control limits will need to be expressed in the same terms as are used on the vertical scale.



Would you say that this process is in statistical control? You might be wise enough to answer that you cannot be sure without inserting the control limits to upgrade the run chart to a control chart. But there was a good reason why I haven't included the control limits here: you might be surprised to know that, starting with the run chart as shown above, the control limits wouldn't fit onto this page! Take a look at the way that Dr Deming drew the control chart of these data with a quite different vertical scale:



Yes, be sure that Dr Deming *did* compute those control limits correctly! So, in order to produce a reasonable picture, he did indeed need to use a drastically different vertical scale from that which would normally have been used when drawing the run chart. Not surprisingly, the effect shown in his control chart is often known as "hugging the Central Line". As soon as Dr Deming saw this "curious condition", as he called it, he knew that there was something for him to find out. What he discovered was well described by his heading to the story: **"Faulty inspection caused by fear"**. His explanation was as follows:

> "The inspector was insecure, in fear. Rumour had it throughout the plant that the manager would close the plant down and sweep it out if the proportion defective on the final audit ever reached 10 per cent on any day. The inspector was protecting the jobs of 300 people."

She was "fiddling the figures"! Could you blame her? Wouldn't you have done the same? Her kind of action is sometimes called "sandbagging". On "good days" she would set aside some of the good items and replace them with some defective ones that she had set aside on a previous day, and on "bad days" she would do the opposite. A similar activity is engaged in by salespeople who need to reach some target in order to obtain a higher rate of commission: I'm sure you can suggest some relevant details! Maybe in some companies the inspector would have been fired when the truth was discovered. But not in this case, I believe. "We reported to the top management our explanation—fear. The problem disappeared when this plant manager migrated to another job, and a new manager came in."

A control chart may not be able to completely solve a problem. But it is certainly a great help for discovering when there is a problem to be solved and can often give some pretty useful clues to aid its solution. In this case the solution was *not* to fire the inspector but instead to kill the rumour.

CONTROL CHART + BRAIN

It is now time to revisit the "simplest guidance" that Dr Deming gave on how to interpret control charts (near the bottom of Day 2 page 15):

> " … if all the points lie between the two control limits then he would judge the process to be in statistical control (stable). Otherwise the indication is that the process may well be out of statistical control, and so then it could be worthwhile to try to identify special causes … "

I also implied that, in practice, it can sometimes be wise not to simply stick to Dr Deming's basic guidance. After all, *he* didn't!

The control chart in the "favourite example" covered in the previous section had no points anywhere *near* the control limits, let alone outside them. Yet, for very good reasons, Dr Deming soon realised that something was wrong. These data were *not* behaving as they would be expected to behave if the process were genuinely in statistical control. Here the points were much too far *inside* the control limits. Remember that the control limits realistically indicate the approximate range of the data to be expected *when all is well*. If all had been well then the control limits would have been much closer together so that, for example, the control chart would have looked more like the charts A1–F1 on page 19. As that was clearly not the case then he knew there was something for him to discover.

This kind of thinking immediately takes us far away from the idea of interpreting control charts by simply and mindlessly following the strict rules of that "basic guidance". Rather than acting as if we were "mindless", let's use our brains. Rather than simply expecting the control chart to tell us all the answers, let's intelligently combine *its* guidance with *our* own judgment and experience. Indeed, that is precisely what Shewhart did when *inventing* the control chart in 1924! If you read the "Discussion on the First Paradox" in the Appendix, do you remember how he began the sentence which contained his famous recommendation for positioning the control limits (at the top of Appendix page 5)? The sentence began with: *"Experience* indicates *…* " *[my italics].*

Using our own judgment and experience along with appropriate guidance is surely what we do with anything important in life. So that is what we need to do when interpreting control charts. It may not be the mathematician's way. But it is the practical way.

Before going any further in that direction, let's focus on a few specific reasons why it makes sense to have some flexibility when interpreting control charts as opposed to "mindlessly" just obeying rigid rules.

1. Just below / Just above

You may recall that the Ford histograms on page 5 showed the positions of Lower and Upper Specification Limits. Deming's disapproval of judging quality merely in terms of conformance to specifications is shown by its inclusion in his list of numerous "Obstacles" (see *DemDim* page 54). This particular Obstacle will be considered in detail on Day 7. But the nub of his argument is this: two things, one of which is just below a specification limit while the other is just above it are, in practice, virtually the same as each other. So what is the logic of saying you must accept one and reject the other? A similar argument most surely applies to points on a control chart. A point which is, say, just above the Upper Control Limit and another which is just below that Upper Control Limit represent measurements which are almost equal to each other. So what is the logic in dogmatically deciding to search for a special cause in the one case but not in the other?

2. Control limits are (almost) never the same again

Further weight to the above argument comes from the fact that, even if a process remains blissfully in statistical control, control limits computed from two different sets of data from that stable process are almost bound to be different from each other. Data vary: thus so will control limits. So if other control limits are computed from different data from that same stable process, the strong likelihood is that the particular two points considered in the previous paragraph will now either be *both* below the new UCL or *both* above it—making it even more illogical to act completely differently depending on which of the two you have.

> **Technical Aid 9**
>
> An important practical point which arises here is that, especially when computing control limits using relatively few data-points (a short baseline), you may occasionally be unlucky with the particular data that you finish up with. Even if the process remains beautifully in statistical control, those data *may* be untypically close together or untypically far apart, with the obvious consequences on the moving ranges and hence on the distance between the control limits. In such circumstances you will hopefully soon become rather uneasy about those limits: before long you will begin to feel that subsequent points are either rather more volatile than you expected or alternatively show the "hugging the Central Line" effect. Don't then be afraid to recompute the limits from a larger or different set of data—it's not "cheating": it's good sense! This kind of situation is not a frequent occurrence—but it *does* happen. This isn't an exact science.

3. Strength of signals

We have already indirectly touched on the matter of "signal-strengths" when comparing Charts A2–F2 with Charts A3–F3 in the "Six Processes" section (page 22). I pointed out there that the signals of the process going out of statistical control were mostly stronger in Charts A3–F3 than in Charts A2–F2; by "stronger" I mean that the points were lying further outside the control limits. The straightforward fact is that, the further outside the control limits a point lies, the stronger is its evidence that a special cause is present: it's not simply a case of either yes or no.

4. Subject-matter knowledge

Obviously, all that the control chart "knows" anything about is how to help us interpret data from a process. It knows nothing else. So, in that sense, it plays the role of the "Un-Knowing" in the Red Beads Experiment. But sometimes, as we saw, the "Un-Knowing" knows just as much as the "All-Knowing" regarding what's important. But not always, of course. Deming was careful to pay due respect to "subject-matter experts". In contrast, I've known teachers of Statistics who seem keen on ignoring subject-matter knowledge lest it "bias" the conclusions available from the data. At the other extreme there are some "subject-matter experts" who perhaps feel they are "All-Knowing" to the extent that they cannot have anything to learn from data! Both extremes are silly. Intelligently *combining* both forms of knowledge makes sense: they *both* have contributions to make. Good subject-matter knowledge might e.g. mean you would need rather stronger signals to convince you to look for a special cause—but not to ignore the data altogether!

> It is interesting to reflect again on the artificial sales data (pages 10–12). It concerned a new product on the market. Surely we would not *expect* that process to be in statistical control in its early days: we would presumably be expecting, or at least hoping for, sales to increase from the starting-point of zero! So it would indeed have been a surprise to see the indication of stability in the control chart on page 11. But, rather than the management team's almost automatic reactions, maybe it would have been wise to heed the control chart's guidance after all. Why might sales keep dropping as soon as the promotions stop? Maybe it *is* a system problem. Maybe the product is of poor quality or too expensive, so that those who

buy it during a promotion do not buy it again and also recommend their friends and colleagues not to buy it either! Thus the problem might, after all, be in the system as the control chart indicates, pointing to the need to change the system by improving the product or charging a more realistic price.

5. Control limits are indications, not boundaries

The control limits indicate the range of values over which you would expect the large majority of data to lie while the process remains in statistical control. But that implies there may also be a small minority of points lying *outside* the control limits while the process remains stable! However, following the above arguments, you wouldn't expect such occasional points to be very *far* outside the limits. A solitary point which "sticks out like a sore thumb" should almost certainly be investigated. It might simply be a mistake —but it might not. Investigation is also likely to be appropriate if you start getting points close to one of the control limits rather more frequently than usual, even if no point actually goes beyond that limit.

6. Other indications

There are other types of indications of special causes that can be seen even without the help of control limits. Examples are apparent trends (as we have seen on page 8) or a special cause might simply shift the process average up or down (as discussed on page 14). Another example, particularly if we are dealing with monthly data, is some kind of seasonal effect. You may have noticed that there was a hint of such an effect in Chart F1 on page 19 (if so, well done!) but, as it was not large enough to send any points outside the control limits, I didn't mention it at the time. I'll comment more on this matter on page 33.

You'll now be appreciating how experience and relevant knowledge of the subject-matter *should* come into play when judging how long particular suspicious-looking behaviours should continue before deciding to look for special causes. It can even depend on the *type* of processes being studied: some processes are more prone to certain special-cause effects than others. For example, deciding whether or not to investigate that hint of a seasonal effect in Chart F1 is surely a matter of judgment along with some process knowledge. However, as we saw in Chart F3, those control limits were adequate enough to detect important special causes. All this and more indicates why on Day 2 I likened the wise use of a control chart to that of judiciously "bending the rules" as stipulated in a tool's basic instruction leaflet (rather than just obeying the "mindless rule") as you become more experienced and knowledgable about using the tool.

The "mindless rule" of course has the advantages of being both easy to express and easy to use. But how can we express the *wiser* approach to interpreting control charts, and what might help us to carry it out in practice? I'll give you answers to both parts of that question that both I and my delegates and students found useful over the years.

Here is my attempt at the first part of the question. Its nature is of course subjective rather than precise—it could not be otherwise:

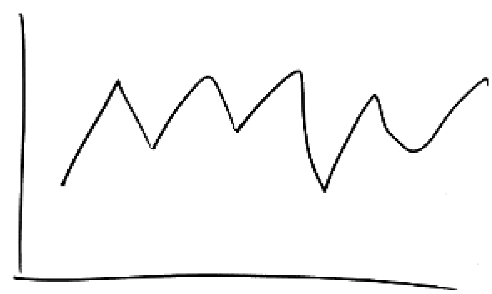
If all (or almost all) of the data-points are comfortably contained within the control limits, and if there are no obvious trends or other patterns visible, then there is no evidence of any special causes affecting the process—so there is no point in wasting time and money looking for any. Otherwise, there may well be.

In the context of the "favourite example" (on pages 24–25), those data-points were of course *far too* comfortably contained between the control limits!

For the second part of the question I'll simply suggest that you repeatedly refer to Charts A1–F1 (or, to be on the safe side, A1–E1). As I've already pointed out, despite originating from a very broad range of processes, those charts all look rather similar (and boring!). Suppose you are ready to interpret a control chart of your own. Does it also look quite similar to Charts A1-E1 (apart from fine detail, of course). Or does it look comparatively *"interesting"*? If so, how? Your answers to those questions will become your guidance for interpreting your chart.

> A contribution to Charts A1–E1 appearing so similar to each other was my choice of vertical scales; they were chosen so that the control limits were similar distances apart throughout. Even this has a precedent from Dr Deming's advice. At one four-day seminar at which I was present, a delegate asked him for guidance on the choice of vertical scale for a control chart. After a brief pause, Dr Deming said "I would suggest one which sets the control limits about two inches apart."

Having just raised the matter of "patterns", I would like to mention something that my good friend Dr Peter Worthington often did near the beginning of some of his seminars. (Peter has almost certainly presented even more seminars on understanding variation than I ever did.) He asked his delegates to sketch what they perceived a run chart of random variation would look like. In his own hand, here on the right is the kind of picture that he told me they almost always produced:



Peter was then able to point out to them that this is definitely *not* random variation! Random variation does not have patterns (except occasionally and briefly by a fluke). This sketch is a *pattern*—it's a *zig-zag* pattern: it is high for a while, then goes down for a while, then goes up again, then goes down again, and so on. Take another look at the in-control charts on the left hand side of page 19—or, for that matter, any of the charts on that page. Do any of them demonstrate such regular up-and-down behaviour? Of course, there are lots of *individual* ups and downs, but they do not occur in a regular and long-running pattern. I repeat that, by definition, random variation does *not* have patterns.

Peter has permitted me to repeat what he told me about how this item in his seminars originated:

> "Interestingly, the idea came from a comment made by an operator in Michelin *[the tyre manufacturer]* in the carbon black plant (awful stuff—it got everywhere including the coffee vending machine!). They had plotted a run chart of the results of an experiment performed on a blend of carbon black and, in passing, I heard the remark: 'That looks random'. 'Interesting', I thought and, being nosey, I took a look—and saw wonderful zig-zagging!"

To watch out for "wonderful zig-zagging" is an important lesson to learn because, as e.g. we shall see this afternoon during part of the Funnel Experiment, a regular zig-zag pattern is a sign of trouble. It most certainly is not what a customer wants to receive from a supplier, whatever the product or service is!

So, in conclusion and following these various arguments, it's clear that intelligent diagnosis of whether or not a process is in statistical control is preferably not just a matter to be decided upon purely by control limits: you need to bring your brain into play as well! The description used several times previously is that a process will be diagnosed as in statistical control if the process data cannot be regarded as different from what could reasonably be expected if the data were being generated by some kind of random mechanism (like the Red Beads or dice again). So notice the important point that concluding a process is in statistical control does *not* imply any "proof" that it *is* a purely random process producing "squeaky-clean" random numbers! (Indeed, one could argue that such a process never exists in practice as opposed to in mathematical theory.) The important consequence arising from concluding that a process is in statistical control is simply that there is no logical way you can identify any special causes from the process-data. In the process-improvement context it follows that, rather than searching for special causes, the available energies and resources should be spent on improving the system—where, after all, there is usually much more to be gained (recall the final Shewhart "bare bone" on Day 1 page 33).

THE SIX PROCESSES REVISITED

On pages 19–23 I described and discussed how our study of the six processes can help us diagnose whether a process is or is not in statistical control. But it can do more than that. If it looks as if the process is being affected by special causes then the control chart can often give us useful guidance about when (and thus often where) to look for them.

So let's conclude this *extra-curricular* discussion by examining more closely than previously the second halves of the process-data. I'll reproduce Charts A3–F3 but now with the points numbered along the horizontal axis in order to aid the discussion. I have already told you what was happening with these processes. But now imagine that I *hadn't* told you and that therefore you *only* had the control charts to guide you. What could they have told you? Also, imagine that you are developing the charts in real time after the halfway stage, i.e. after having computed the control limits from the first 24 data. Since at that stage the charts indicate the processes to be in statistical control, it is sensible to extend those control limits into the future and add the next 24 points to the chart one at a time as and when they are obtained.

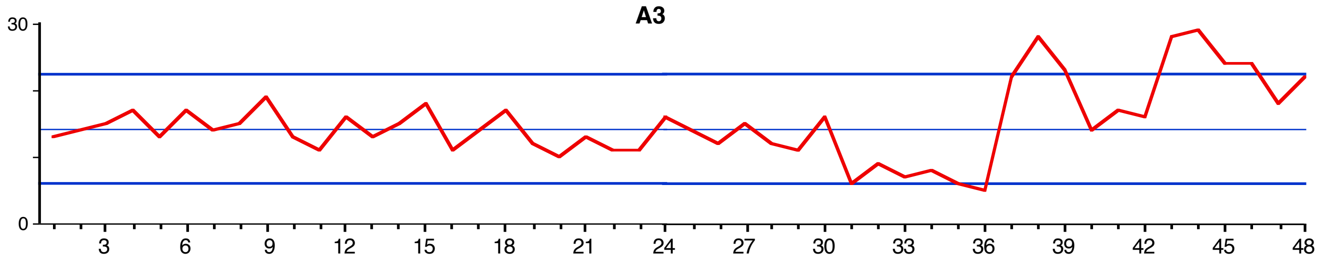
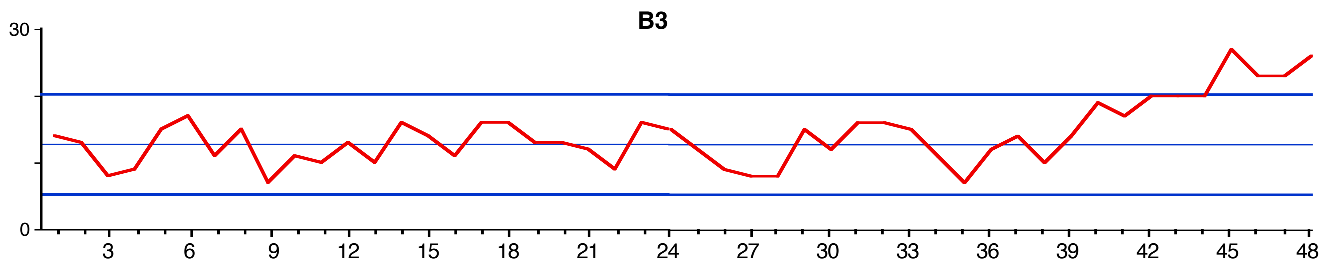


Chart A3 is clearly indicating stability up to and including Point 30, but then Point 31 suddenly drops onto the LCL (Lower Control Limit). As recently pointed out, such a point may well occur occasionally even if the process remains stable. However, after seeing the next one or two points again lying near the LCL and also clearly lower than any of the first 30 points, there is little doubt that a special cause has occurred which has lowered the process average. It would thus be sensible to try to identify the special cause which occurred between Points 30 and 31. Subsequently perhaps some remedial action was taken after point 36 which however soon appears to have rather *over*compensated for the drop! From there until the end of the chart all but one of the points are above the Central Line, with several points above the UCL (Upper Control Limit). In fact, the first three of these points are all near or above the UCL, and this is already very strong evidence that the process average has suddenly moved higher than it was in the initial stable period.

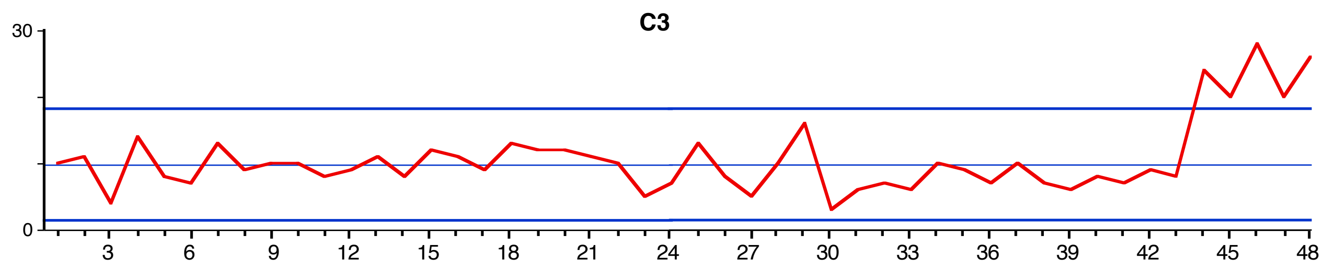
You will see that this interpretation of the chart accurately reflects what we recall was the truth: four dice were used for Points 1–30, two dice for Points 31–36, and six dice for Points 37–48.



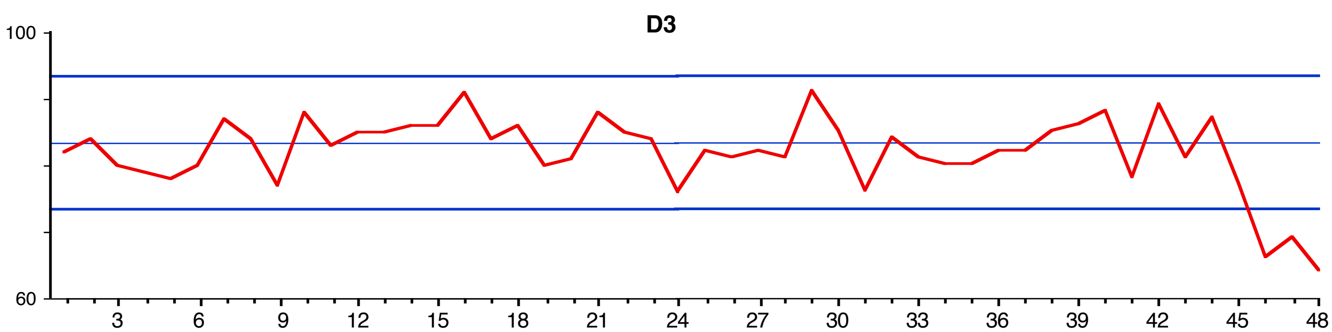
Nothing "interesting" seems to be happening in Chart B3 until around Point 40 which is close to the UCL. As in Chart A3, by itself this is only a tiny hint that something untoward may be happening, but that hint immediately gets supported by point 41 (also close to the UCL) and then confirmed by Point 42 (virtually on the UCL). By coincidence, the following two points are both the same as Point 42, but we hardly need them to convince us that the process has moved upward. The remaining points are even higher and so we appear to have a trend rather than just a sudden move (as occurred twice in Chart A3).

That again is an accurate interpretation of what actually happened: two extra coins were added every time from Point 39 onward. This was a relatively steep trend and so was spotted soon after it began. Had the slope been gentler then it would have taken a little longer to become confident that it was happening; however, at that stage one would still be able to trace back to see roughly where the trend began, which would help identification of the special cause that was producing it.

> With regard to these illustrations using dice, it may be worth my reproducing a suggestion from *EST* page 50: "Such 'games' *[as in these illustrations]* are a fast and effective way to gain experience of constructing and interpreting *[control]* charts. Work with a colleague. One of you generates the data, now and again unobtrusively changing the process. The other records the data and draws and interprets the charts. You could also try not bothering with the *[control]* limits, and see how you get on!"



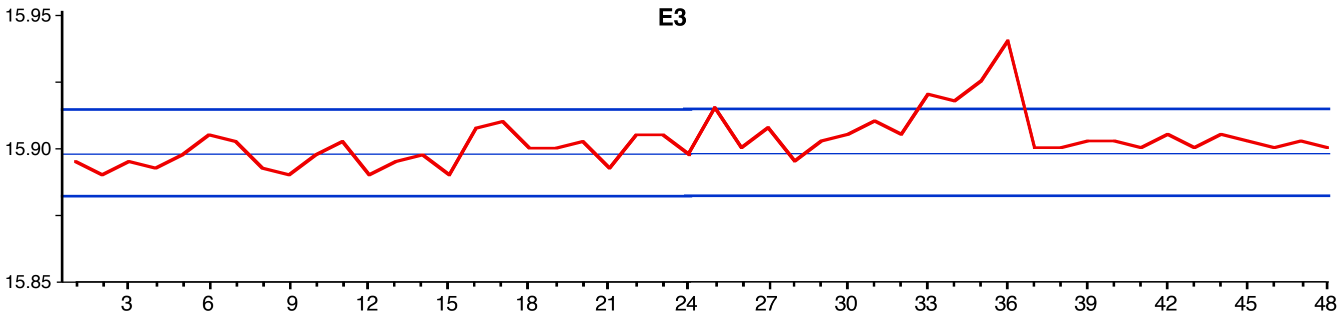
There is little to discuss with the Red Beads illustration: you will recall that I contrived to double the recorded values for the last six points. In fact, it looks as if I only doubled the values for the final *five* points; however, the first actual count of red beads when I started doing this happened to be only 4, so Point 43 is at 8—thus appearing to be just *before* the special cause occurred. You can't expect the chart to be absolutely right all of the time!



The control chart of my morning pulse rates shows stability up to Point 44. That was the day when my doctor prescribed the beta-blocker and I took the first tablet straightaway. My pulse rates on Days 44, 45 and 46 were respectively 87, 77 and 66 (you can't really see the 77 on the chart as it's in the middle of an almost straight line). As you might guess just from the final three points on the chart, the pulse rate then settled down to stability at a much more healthy level! Obviously there is no doubt here as to what the special cause was, but I confess I was both amazed and delighted by how fast and how effectively it worked!

> This raises a rather obvious question which fortunately also has a rather obvious answer. Regarding the control chart, what do we do now? This was a *good* special cause. It has not *harmed* the behaviour of

the process as do most special causes: it has improved it. So, of course, we do not try to remove a good special cause: we retain it and, if possible, incorporate it into the system (of course, easy enough in this particular case). But since it has changed the system, clearly the current control limits are now redundant. Thus, after a few more data have been recorded, the obvious thing to do is recompute the control limits using data from the changed system and, presuming the changed system is seen to be in statistical control, to extend those new limits into the future.



The data in Chart E3 are taken from an ancient but fascinating case study which Don Wheeler relates in a chapter on "Using Process Behaviour Charts for Continual Improvement" in his book with David Chambers: *Understanding Statistical Process Control*[b]. This is how Don introduces the story:

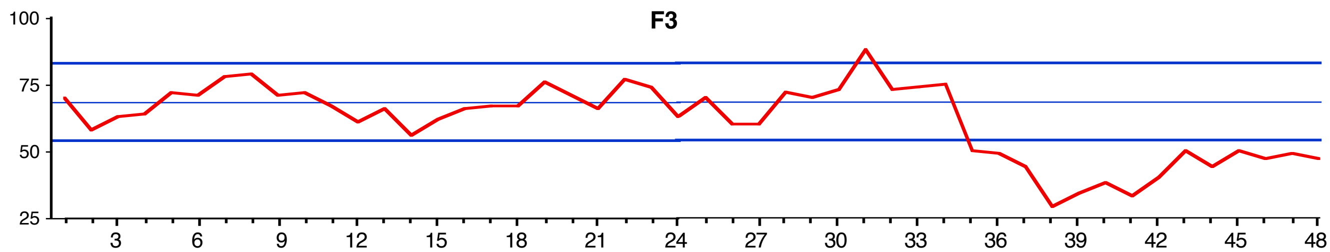
> "The … chart was brought to this country by a group of executives from the Body and Assembly Division of Ford Motor Company, following a visit to the Tokai Rika Company in March, 1982. As the Ford group was touring the Tokai Rika plant, they observed eight production workers 'engaged in active discussion' around this ... Chart. To the people from Ford, it seemed that something must be wrong with the process represented by the chart, so they asked about it. They expected there to be an internal production problem, or an assembly plant problem, or a problem of too many rejects. However, they were told that this was simply a routine review of an ongoing process and, in fact, the process was currently operating predictably and was well within the specifications. To substantiate this, their hosts translated the chart, and presented a copy to the Ford group … "

The control chart (all hand-drawn and several feet long!) covered the period from August 1980 to March 1982. The points in Chart E3 are the daily averages of four measurements of the distance between the flange and the detent in manufactured cigarette lighter sockets. The nominal value of this distance was 15.90 mm with specification limits 15.80 mm and 16.00 mm. (So take a look at the vertical scale on Chart E3—doing well, weren't they?)

The high point in Chart E3 (Point 36) was for Friday 26 September 1980, and Points 37–41 were for the following week: Monday 29 September to Friday 3 October. Using the notes in the translation, Don writes:

> " … 'abrasion on the positioning collar' is identified as the Assignable *[Special]* Cause for the process excursion noted in late September, 1980 ... in addition to writing down the Assignable Cause on the chart they also took action—the very next day the process average shifted back to the target of 15.90 mm. Again, a note on the chart tells what was done.
>
> As a temporary solution, a worker turned the worn collar over to use the back side. Two days later a new collar was installed. This incident displays a desire on the part of the Tokai Rika personnel to operate at the target. The process was in no danger of producing nonconforming product, yet they took the trouble to fix it so that it would stay centred on the target value of 15.90 mm. Moreover, just as the shift on September 29 shows the desire of the workers to operate at the target value, the replacement of the collar on October 1 shows the support of the management for this policy."

It will probably not have escaped your notice that, as soon as action was taken to prevent further deterioration due to the worn collar, the variability of the data became considerably *less* than it was in the first half of the chart. The reason for this was not noted on the chart. Maybe the new collar was of an improved design or made of better material. Maybe the old collar had already been causing some deterioration earlier on but not enough to produce out-of-control signals on the control chart. Whatever was the case, the Tokai Rika personnel did the obvious thing. Yes, they recognised the existence of the changed system by recomputing the control limits and extending those limits into the future. Notice how wholly irrelevant any consideration of "conformance to specifications" was to how they behaved. This account about the Tokai Rika chart is considerably expanded upon in Part B of the Optional Extras.



Finally, recall that the data in Chart F3 show the monthly US trade deficits for the years 2006–2009. The striking feature is the considerable downward trend following the peak value at Point 31, particularly the drop to below the LCL at Point 35, and then the continuing sequence of points below the LCL until the end of the chart. But what was Point 35? It was November 2008, the first full month after the seriousness of the global financial crisis became clearly evident. Naturally, this caused internal demand to plummet, thus considerably reducing imports. Export figures also fell, but not to the same extent since several of America's overseas markets were less seriously affected by the crisis.

Let's return to the interesting feature observed on page 28 that, rather than being properly in statistical control, there was some indication of a seasonal effect during the first two years. Examining the detailed figures over several years, this effect is seen to be real and primarily comes from rather higher and lower import figures in the summer and winter respectively. One could therefore consider switching to using "deseasonalised" figures. However, this seasonal effect is not particularly large and is relatively easy to interpret, so there is little harm in continuing to use the raw data. In fact, this is also consistent with something that Shewhart taught. He was keen, whenever possible, to use understandable raw data rather than data which have been through some mathematical manipulations that might make them "tidier" in some way but less easy to see what they actually represent.

As mentioned on page 21, the Springboard article shows the earlier version of Charts F1–F3. It is interesting to compare the two versions. For example, the earlier version of Chart F1 hardly hints at the seasonal effect indicated in the later version. As another example, the two versions of Chart F2 look remarkably similar, both showing the considerable decline at the crisis time followed by the partial recovery. But in the earlier case the recovery is shown to be moving back into at least the lower reaches of the pre-crisis system whereas, in this later version, the values are beginning to settle down into a region distinctly lower than the previous Lower Control Limit—verifying the greater seriousness of the 2008 crisis compared with the relatively minor recession in 1990.

Lastly, although we have been using Charts A3–F3 in this discussion, do not forget the importance of Charts A2–F2. These showed control limits which had been computed from the *out-of-control* data in those charts, so that the limits suffered some "contamination" compared with the limits used in Charts A3–F3. But recall that, as pointed out on page 22, despite that contamination, pretty much the same conclusions would have been reached with Charts A2–F2 (and at the same times) although the signals are usually not quite as strong. It is well worth re-emphasising the incredibly valuable feature that control charts can often do a great job *even if* their limits are computed when the process is out of statistical control.

As you can see, all we have left on this page is a couple of Technical Aids. So if you are on Stats-level 0 then please move straight on to page 35.

Technical Aid 10

In the way that I set up Charts A3–F3, the control limits had been computed while the processes were in reasonable statistical control: thus it made sense to extend those control limits into the future. But suppose that had not been the case. Consider, for example, my control chart on page 18. Suppose I had just reached the halfway point where I had available the first 12 data-values and had just computed the control limits and drawn the control chart over those 12 points. Now, although none of those 12 points were outside the limits, the indication of a trend was already very strong. So what would I have done?

I might already be so convinced of the trend that I would start searching for its special cause straightaway. Or I might have waited until I had seen two or three further points for confirmation. The fact is that when there is immediate evidence (i.e. as soon as the control limits have been obtained) of the process being out of statistical control then there is, of course, no sense in extending those control limits into the future. The process is not predictable—so there is no sense indicating on the chart what the prediction *would* have been if the process *had been* predictable!

Thus (presuming the process is one over which you have some influence) your need is to try to identify and deal with the special cause(s) of whose presence you have now been made aware. After you have taken appropriate action to try to stabilise the process, then (as in the Tokai Rika case study, Process E) you would resume recording data and, in due course, recompute the control limits. You are thus "back to Square 1": extending those control limits into the future if the process is now in statistical control or else resuming your search for special causes if it isn't.

Notice that if, as in the Tokai Rika case study, the action you have taken has not only removed the special cause but has also reduced the common-cause variation, it is quite possible that further special causes may become visible on the chart. The narrower control limits may now reveal special causes that were already there but had been camouflaged by the previous greater amount of common-cause variation.

Technical Aid 11

Finally, here are four "tidying-up" points.

(a) Besides the Red Beads Experiment (Process C), Process B (counting the number of Heads when 25 coins were tossed) also fits the "batch inspection" conditions (see page 13). So the method used for computing the control limits for Process B was the same as for the Red Beads Experiment except with n = 25 instead of 50. The moving-range method was used for the other four processes.

(b) The control limits for all six processes were computed from 24 data. As mentioned earlier, this number was chosen simply because that was what we were used to using for the Red Beads Experiment. However, remember that for "live" charts a shorter baseline is recommended (Technical Aid 8 on page 17).

(c) If you refer to Dr Wheeler's own account of the Tokai Rika case study, you may notice that the control limits used in Charts E1 and E3 are different from the original. This is because the control limits used here were computed from the first 24 data-points on the Tokai Rika chart given to the Ford personnel. It is unclear from the account when Tokai Rika's control limits were computed, but it was probably before the beginning of the chart as we have it.

(d) There is a further illustration of the moving-range method on page 9 of the file "Q. Contributions from Balaji Reddie".

INTRODUCTION TO THE FUNNEL EXPERIMENT

One more time. Dr Deming warned us that

> "if we wade in at *[problems]* without understanding, we only make things worse. This is easy to prove."

As with the Experiment on Red Beads, the Funnel Experiment may just look like a rather silly game when you are introduced to it. Hopefully, like the Red Beads Experiment, I think that before long you'll realise it's rather more than that.

As we know, the Funnel Experiment was created by Lloyd Nelson and subsequently described and discussed by Dr Deming in his four-day seminars. But, unlike the Red Beads Experiment, it was not actually *carried out* in the four-day seminars. There was a good reason for that. In the form suggested by Dr Nelson, it is only feasible to carry it out for real either individually or in small groups—not the several hundred at a four-day seminar! In that original form, you need:

- a funnel,
- a marble which is small enough to pass through the funnel,
- a table, exactly horizontal and covered by a thick tablecloth …
- … on which is marked a target-point—this is where (in our silly game!) we would ideally like the marble to come to rest after being dropped through the funnel and bouncing or rolling around on the tablecloth,
- a portable stand on which to fix the funnel at some convenient distance above the table and which can then easily be moved to different positions on the table,
- a ruler to measure how far the marble is off-target and a protractor to measure angles,
- —oh, and permission to record the resting-places of the marble on the tablecloth with a marker-pen!

Believe me—my version is easier to perform and also results in less laundry costs! Basically, it's a one-dimensional version of the experiment, whereas the Funnel Experiment itself is two-dimensional. I.e., this is a "flattened" version of the Funnel Experiment where both the funnel and the marble can only move either to the left or to the right (one dimension) rather than anywhere horizontally (two dimensions) in the original version. A further advantage is that, unlike with the original version, we will now be able to analyse data produced in the experiment using tools we already know about: histograms, run charts and control charts. That would be, to put in mildly, rather more difficult with two-dimensional data!

> There will be four strategies to examine during the Funnel Experiment. To save time I shall just ask you to construct histograms for the first two strategies and run charts for the other two. I shall return to the Funnel Experiment in Part A of the Optional Extras section where I will also show control charts in operation on data from the experiment.

So, in our one-dimensional version, imagine we have *two* tracks both like the following:

with one track suspended some distance directly above the other. (You will notice that the central number 30 is coloured differently from the others: 30 is the "target-point" in our silly game.) The funnel's track is the higher one: it has a hole in the middle of each square so that the funnel can be suspended at any numbered position. The marble's track is the lower one: the sides of the numbered squares are rigid but the insides of the squares are made of some elastic material. So the general set-up is that (a) the funnel can be placed at any position on the higher track and (b) the marble can then be dropped through it onto the lower track where it will briefly bounce around and then finish up in one of its squares—possibly directly under the funnel or, more usually, displaced by one or more squares to the left or right. To complete the image we should join the two sides of the top and bottom tracks by vertical sheets of glass or transparent plastic to prevent the marble bouncing off its track but so that we can still observe what's happening!

However, you don't have to build that whole model! During the description I shall show the positions of the funnel and marble along their tracks by little icons on a *single* track. So, for example, if the funnel is at position 31 and, after being dropped through the funnel, the marble bounces around and finishes up at position 33 then I shall show this as:

So how exactly will you be able to carry this out with your less extravagant set-up (of which details follow)?

In our timing schedule we are now coming up to the lunchtime break. If you have not already done so, you will now need to prepare your own track like the one illustrated in order for you to be able to carry out our version of the Funnel Experiment, ready to use when we resume this afternoon. On page 59 you will find a larger copy of the track in two parts, one part going from 20 to 30 and the other from 30 to 40. Preferably (but only if convenient) I suggest that you copy or paste these (either electronically or using glue!) onto some card. Then join them together into a single strip as shown above by attaching the two 30s one on top of the other. I've also included on page 59 larger images of both the funnel and the marble: you could similarly put the image of each that you choose onto small pieces of card to use in the experiment. Alternatively, you could e.g. use a large silver coin to represent the funnel and a small copper/bronze coin to represent the marble.

Finally, as previously warned (on Day 2 page 44), you will also need two standard dice (and a shaker if you so wish!).

Then, fully equipped, you will be ready to carry out our version of the Funnel Experiment. When you begin, I recommend that you work through the first few pages of the Major Activity quite slowly and carefully since they will lay the foundations of how you will be spending most of the time this afternoon.

NB There will be no separation between Stats-level 0 and Stats-levels 1–3 this afternoon, and so then the timing guidance will only be included on the right-hand side rather than on both sides as during this morning.

If you are particularly numerically inclined, you *could* conceivably carry out this experiment without using the track: you might instead be able to *visualise* what is happening on the track by just interpreting the numbers which will develop. However, I believe most people will find it easier to actually *see* what is going on, at least in the early stages of operating the various strategies to be studied during the experiment.

But once you have become familiar with any particular strategy then you may soon find you can carry on without continuing to use the track. Use it for as long as you personally find it useful to do so and then, if you prefer, don't bother with it any more. However, since I have worked through this entire Major Activity many times, you may find my own experience to be useful here. I soon found that it was definitely safer to use my track throughout *every* stage of the first strategy, since it is actually the trickiest of the four to carry out. I also worked with the track throughout the third strategy before realising that there was a useful and simple short-cut (which I'll show you). I used the track for a little while with the fourth strategy but soon found that I didn't need it any more. The second strategy is the simplest of all, and there I hardly needed the track at all. I shall reflect the relative ease or difficulty of the various strategies in my timing guidance.

Your version of the track going from 20 to 40 will be sufficient for much of the afternoon. However, especially later on, you may occasionally need extensions outside that range. If so then you will find them on page 61.

NB Just In case you really cannot find any dice to use then I have provided below a couple of sample sequences of dice-throws for you to choose between and use instead. Note that both of these sequences begin at Stage 6: I'll be providing the first five stages for you this afternoon when I shall use them for demonstration purposes.

Stage	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
Seq. A	6,4	2,4	4,5	4,5	1,3	5,6	6,3	3,5	6,4	3,5	2,2	5,2	4,2	2,4	6,6	4,6	5,3	4,3
Seq. B	3,5	3,6	4,4	1,3	1,1	4,1	5,3	1,6	4,5	1,2	5,4	1,4	6,1	1,2	1,3	5,4	6,5	6,1

Stage	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Seq. A	2,5	6,3	5,6	6,3	5,1	6,1	2,1	2,2	3,6	5,3	6,6	5,3	4,6	6,5	3,4	3,6	5,6
Seq. B	4,3	6,6	1,5	6,3	4,4	5,2	4,4	1,4	6,1	3,2	5,5	5,6	6,2	1,3	4,6	6,3	3,1

Relevent sections of this Major Activity are also in the Workbook: I'll give you the page references where and when appropriate.

MAJOR ACTIVITY 3-h

Before we begin, a few words about the nature and purpose of the Funnel Experiment, some of which you will recognise as having been just as relevant to the Red Beads Experiment. As there, sooner or later it will become obvious to you that some of the strategies used and actions carried out are, depending on the circumstances, rather foolish! But that is precisely what's intended. Both experiments are *very* simple—so simple that the difference between good and bad practice becomes plain for all to see. However, after the Funnel Experiment had been presented, both Dr Deming in his enormous seminars and I in my smaller-scale ones would then ask the delegates to tell us of examples of what had now become clear to them as costly and damaging bad practices in their own lives, their own experience, their own workplaces, their own organisations. And invariably the delegates would come up with *dozens* of such real-world examples analogous to the bad practices of which they were now aware through what they had just learned from the Funnel Experiment. Often they had carried out such practices themselves—not because they were bad people but because they hadn't previously *realised* that the practices were so bad. Now they did. Thus, as with the Red Beads Experiment, the purpose of the Funnel Experiment is to alert you to the truth of such matters, so that in future you will be able to figure out more successfully what is good practice and what is not: what to do and what to avoid. As you go through the experiment, when similarly you think of practical examples in your own experience of what the Funnel Experiment is teaching, make a note of them since they will be of help when you get to today's final Activity.

It is easy to think of real-life illustrations of the Funnel Experiment in many contexts. For example, consider my school bus again. Suppose the bus driver (who actually was my uncle!) was really keen to arrive at my bus-stop at exactly 8.30 am—which thus became his *target*. How might he use past experience to help him do that? Or we could think in terms of a manufacturing process. Suppose the nominal width of a socket, made by injection moulding, is 2.30 cm. How might one use information about past measurements of such sockets to try to manufacture future sockets closer to the nominal value?

With reference to the two illustrations just mentioned, the numbers on the funnel's and marble's tracks can relate to *minutes* in the case of the bus arrival time and *tenths of a millimetre* for the width of the socket. So the "target" of either 8.30 am or 2.30 cm is represented by 30 on the track, while the 29 refers to 8.29 am or 2.29 cm, 31 to 8.31 am or 2.31 cm, 34 to 8.34 am or 2.34 cm, etc.

The First Two Rules of the Funnel

An obvious way to attempt to improve results from a process is to compare the current outcome (today's bus arrival time or the width of the socket just manufactured) with the nominal or target value, a comparison which may suggest an adjustment to try to improve the next outcome. In our "flattened" version of the Funnel Experiment, this adjustment will be represented in what follows by the movement of our funnel along its track—to the *left* if we want to try to make the next outcome respectively *earlier* or *smaller*, or to the *right* if we want to make it *later* or *larger*. Thus, if the bus arrives today at 8.33 (three minutes *late*), perhaps the bus driver should set out three minutes *earlier* tomorrow. (I am assuming that my uncle is keen to please *me* rather than anybody else!) Or, if the socket just made measures 2.33 cm (0.03 cm too *wide*), we can adjust a control on the machine to *reduce* the average diameter of future sockets by 0.03 cm. In either case, the equivalent movement of our funnel is a distance of 3 squares (minutes or tenths of a millimetre, etc—whatever units we are using) to the *left* along its track. Thus we *compensate* for the amount that the current outcome is off-target by moving the funnel by that same amount in the *opposite* direction. This type of compensation strategy is in fact the *second* of what both Drs Nelson and Deming referred to as the four Rules of the Funnel. Here we are demonstrating it first because it corresponds to what was happening first in the Ford example on page 5. We shall introduce the other three Rules in due course.

Whatever is tried, the fact remains that different things (be they bus journeys, manufactured sockets, or anything else) are almost always different! This is because, as we are now well aware, all processes have their *common-cause* variation as well as possibly some additional special causes of variation.

With the equipment as described, when the marble is dropped through the funnel it will finish up either to the left or right of the funnel or sometimes directly underneath it. There is no need to use anything very complicated to model the variation between the funnel's current position and where the marble finishes up. Anything reasonable will produce the main messages to be learned from the experiment. So that's why you have your two dice. You'll throw the dice and add up the two numbers showing: let's call that the *dice-score*. Since the faces on the dice range from 1 to 6, obviously you will get a dice-score of between 2 and 12. Then use the dice-score *in conjunction with the following little table* to decide where the marble finishes up relative to the funnel's current position. Thus e.g. if you throw two ones, so that your dice-score is only 2, you place the marble 5 (= 7 – 2) squares to the left of the funnel's current position. Or if you throw a 3 and a 5, with dice-score 8, then you place the marble one (= 8 – 7) square to the right of the funnel.

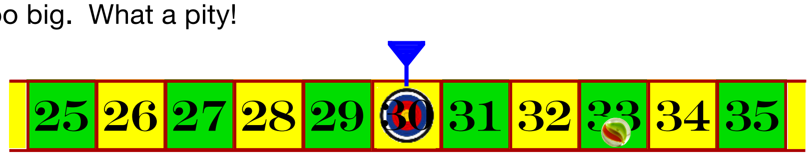
Dice-score	2	3	4	5	6	7	8	9	10	11	12
relative to	5 left	4 left	3 left	2 left	1 left	Under	1 right	2 right	3 right	4 right	5 right

From now on I shall use the colour-coding in this little table to help you find your way around: the funnel in blue, and the dice-score in brown. The marble's position (and thus the outcome or result) will be in red.

In this example, since of course the preferred outcome is always that target of 30 (corresponding to the bus arriving at 8.30 or to a socket diameter of 2.30 cm), you may as well start by putting the funnel at what would appear to be the "obvious" position, i.e. 30:

You throw your dice. If you are fortunate enough to get a dice-score of 7 (e.g. by throwing a 2 and a 5), the table shows that the marble finishes up directly under the funnel—dead on target!

But suppose you're not that lucky. Perhaps your dice fall as a 6 and a 4, giving you a dice-score of 10. Then the table shows that the marble finishes up 3 squares to the right of the funnel, i.e. 3 to the right of 30 giving 33—that's 3 too big. What a pity!



So let's try to make the next outcome a little smaller, to attempt to get it closer to that target value of 30. That was indeed the "logic" behind what was happening at Ford. As discussed earlier, according to Rule 2 (which, recall, was Ford's First Strategy) the bus driver would then leave the depot 3 minutes earlier tomorrow, or we would adjust the control on the injection moulding machine down by 0.03 cm. So equivalently you move the funnel by 3 squares to the left along its track, i.e. from 30 to 27. (From now on I shall not keep referring directly to those two illustrations, as otherwise the description will become very lengthy.)

A particularly appealing way of considering this strategy (i.e. Rule 2 of the Funnel) is that it tells you to move the funnel to the position (27) where, *if only* it had been there when the marble was just dropped through it, you *would* have just obtained the preferred target outcome of 30! With the funnel placed there, the dice-score of 10 led you to move the marble 3 squares to the right of the funnel; so you *would* then have had:

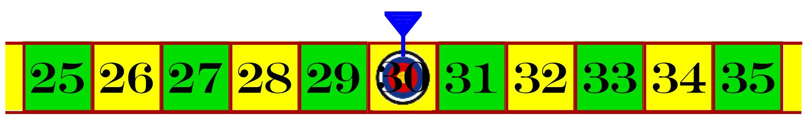
Now, in order to record the progress of Rule 2 in an organised fashion, both to study its behaviour and, later on, to compare it with the other three Rules of the Funnel, you will need to carry out some systematic book-keeping! So let me get you started with your own track and whatever you are using for your funnel and marble. I'll use our familiar symbols for the funnel and the marble, and also for the target (30 in our illustrations).

We'll use my own dice-scores to start with so that you can follow my illustration exactly: you'll be able to start throwing your own dice very soon! *Work through these demonstration stages carefully on your own track so that you can be sure about what to do when I leave you on your own.* Take it steadily at first—remember (as my mother used to tell me!): "more haste, less speed".

The first stage as just described was as follows:

To start with, ...
... as suggested, put the funnel at 30

Stage number 1
starts at **30**

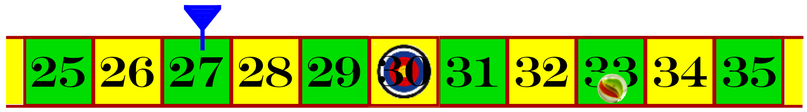


Suppose your first throw of the two dice produces 6 and 4
From the little table, this means the marble finishes up:
So the marble finishes up at 30 + 3 = 33

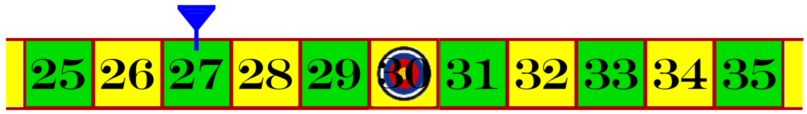
Dice-score = 10
3 to the right of
is at **33**

Since 33 is 3 to the *right* of the target 30 …
… you move the funnel by the same amount *in the opposite direction*
Thus the funnel is now at 30 – 3 = 27

Since is 3 to the *right* of ,
move 3 to its *left*
is at **27**



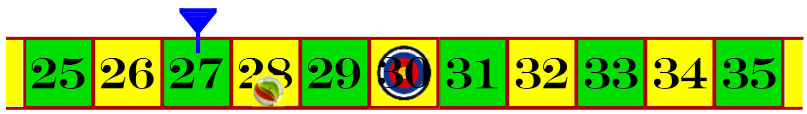
At Stage 2, it's the same operation except with the funnel starting at its new position of 27.



You throw the dice again, and suppose this time you get a 3 and a 5.

You are now at ...
The funnel is at 27
Your second throw of the two dice produces 3 and 5
From the little table, this means you put the marble:
So the marble finishes up at 27 + 1 = 28

Stage number 2
is at **27**
Dice-score = 8
1 to the right of
is at **28**



Since 28 is 2 to the *left* of the target 30 …
… you move the funnel by the same amount *in the opposite direction*
Thus the funnel is now at 27 + 2 = 29

Since is 2 to the *left* of ,
move 2 to its *right*
is at **29**

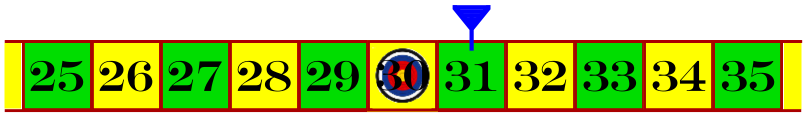
So here we are at the start of Stage 3:

In Stage 3, suppose your throw of the dice produces two 3s. Using your marble and funnel on your track, check that you agree with the following details:

is at **29**
Dice-score = 6
goes 1 to the left of
is at **28**

Ah, but that's exactly where Stage 2 had left it! The fact that the marble is at 28 again is, of course, a coincidence. The marble is there this time for entirely different reasons from the previous time: different position of the funnel and different dice-score.

Next, since is again 2 to the left of ,
again move 2 to its *right*
is at **31**



At Stage 4 suppose your two dice produce 1 + 6 = dice-score 7 and at Stage 5 they produce 3 + 1 = dice-score 4. Working as before, the book-keeping so far can be summarised as follows. To save space I am now abbreviating e.g. "3 to the right" by 3R and also indicating "marble directly underneath the funnel" by ⇓. Now, using your track, carefully check through all these initial stages. Start again from the beginning to get into the swing of it! Then continue through the two new stages.

Stage number	1	2	3	4	5	6
is at	**30**	**27**	**29**	**31**	**30**	**33**
Dice-score =	10	8	6	7	4	
From , then goes	3R	1R	1L	⇓	3L	
Outcome: is at	**33**	**28**	**28**	**31**	**27**	
As relative to is ...	3R	2L	2L	1R	3L	
... you move the funnel	3L	2R	2R	1L	3R	
So is now at	**27**	**29**	**31**	**30**	**33**	

One case which hasn't occurred during these first five stages is where the marble finishes at 30, i.e. exactly on-target. In that case, Rule 2 simply leaves the funnel where it is. So when this happens I suggest you write √ in the "As relative to is ..." row to indicate that the marble is bang on target, and also put a dash — in the "... you move the funnel" row to indicate that then you *don't* move it!

OK, over to you—throw your own dice from now on! Continue the experiment through a total of up to 40 stages. Have patience—as mentioned earlier, this is by far the trickiest of all the four strategies to carry out: you'll be able to work through the other three much more quickly. (Keep an eye on my timings though: if you get short of time then be content with fewer stages.) For your convenience, I have reproduced below the little table which shows you how to interpret your dice-scores.

Dice-score	2	3	4	5	6	7	8	9	10	11	12
relative to	5 left	4 left	3 left	2 left	1 left	Under	1 right	2 right	3 right	4 right	5 right

The table in which to develop your data is on page 44 *[WB 38–39]*. On page 45 I shall then help you to construct your histogram. If you are using the Workbook, after completing the table there then return to page 45 here for advice on forming your histogram before returning to Workbook page 39.

NB Looking ahead, you will see on page 46 that I shall advise you to use the *same* sequence of dice-scores in all four parts of the experiment. This is both to save time and also because it is interesting to see how the four Rules of the Funnel produce noticeably different behaviours *using that same* sequence of dice-scores—in particular, you will then know that the different behaviours *do* result from the differences between the Rules rather than from different sequences of dice-scores.

Of course, to carry out that advice could involve the annoyance of a lot of page-turning! So, to avoid that, I suggest it would make sense (if convenient) for you to take a copy of the first table as soon as you have completed it. Alternatively, again if convenient, you could simply temporally detach the table from your binder. In case *neither* of these alternatives is convenient, you will still be able to avoid the repeated page-turning by carrying out a little manual copying from the table. I'll show you the best way to do this in the paragraph following the table on page 46.

> **Suggestion:** If you are working in a small group rather than on your own, I suggest that each of you produce your own data. You can learn a great deal by comparing your different sets of results to discover which features are similar to each other and which are not. However, particularly if you are studying on your own, there are two runs of the whole experiment summarised in the Appendix for you to compare with your own results. Do your own experiment first though!

So off you go to Stage 6 on page 44 *[WB 38]*.

This table is also on Workbook pages 38–39.

Stage number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
is at	**30**	**27**	**29**	**31**	**30**	**33**								
Dice-score =	10	8	6	7	4									
From , then goes	3R	1R	1L	⇓	3L									
Outcome: is at	**33**	**28**	**28**	**31**	**27**									
As relative to is ...	3R	2L	2L	1R	3L									
… you move the funnel	3L	2R	2R	1L	3R									
So is now at	**27**	**29**	**31**	**30**	**33**									

Stage number	15	16	17	18	19	20	21	22	23	24	25	26	27
is at													
Dice-score =													
From , then goes													
Outcome: is at													
As relative to is ...													
… you move the funnel													
So is now at													

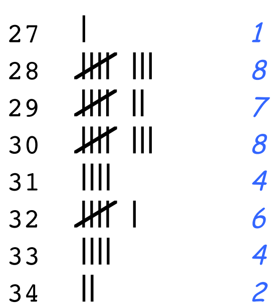
Stage number	28	29	30	31	32	33	34	35	36	37	38	39	40
is at													
Dice-score =													
From , then goes													
Outcome: is at													
As relative to is ...													
… you move the funnel													
So is now at													

Finally, summarise the outcomes that you've just generated in a histogram. The "outcomes" are the resting positions of the marble, i.e. the positions of  in the yellow-shaded rows in your table. However, now that you have as many as 40 data-values to include in the histogram, we need to think more carefully than before about what would be an efficient way to produce it.

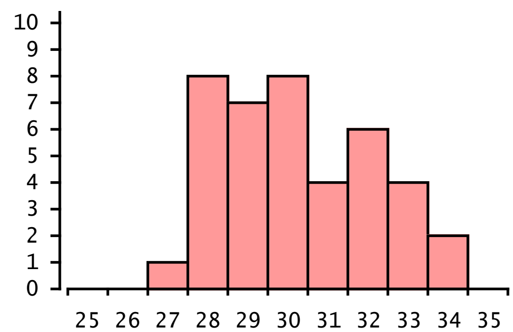
To construct a histogram, most people find it effective to first produce a "tally chart". Here's how. With data such as those you have just produced, write down all the different values you've obtained as a column (see on the left below). Then trace through your data and at each value write a mark | against the appropriate number in the tally chart. So, after you've seen a particular number four times, you will have written |||| against it. But if and when you reach the fifth occurrence of that number you instead draw a line through the previous four marks: ~~||||~~. For future occurrences of the same number you again write single marks until you reach the tenth, at which time you do the same as at the fifth; so you'll then have written ~~||||~~ ~~||||~~ against that number. After you've completed doing this it will then be quick and easy to add up the totals and hence draw the histogram.

So suppose that, after you've been through your 40 values, you've finished up with the tally chart shown below. It will then take very little time to produce the totals (in *blue*) and hence construct the histogram as shown on the right. Be reasonably neat, but there's no need to be too precise—freehand will do. No need to find a ruler!

Tally chart



Histogram

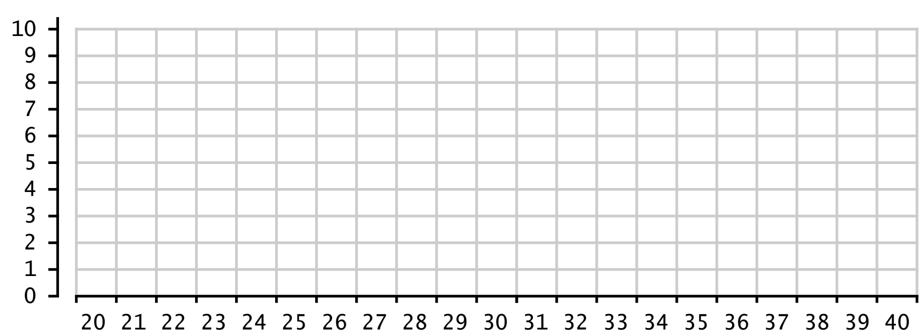


So now please carry out this same operation using your own values of  (including my first five values) from the yellow-shaded row in your book-keeping table on page 44 *[WB 38–39]*.

A larger version of this space for the tally chart and histogram is on Workbook page 39.

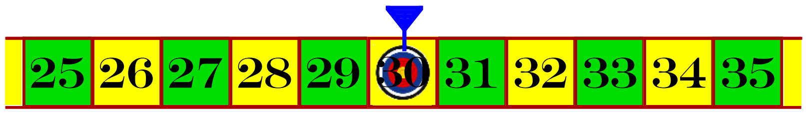
Tally chart

Histogram



I'll postpone discussion on the outcomes from this strategy (which you'll recall is what Drs Nelson and Deming called Rule 2 of the Funnel) until you've repeated the whole exercise using an easier strategy.

In contrast to Rule 2, let's now be relatively idle by simply putting the funnel at 30 and *leaving it there*, *irrespective* of the outcomes, i.e. of where the marble finishes up. This strategy, which is Rule 1 of the Funnel, is of course equivalent to the Ford personnel switching off their automatic compensation device (which was their Second Strategy). So you'll have some idea about what to expect in what follows!



As suggested on page 42, I recommend that you don't bother to throw your dice any more but just use the same dice-scores as you obtained the first time. I mentioned two advantages of doing this on page 42.

In this case where you don't move the funnel at all, you may be able to stop using the track almost straightaway. As I've just said, you simply put the funnel at 30 and *leave it there*. Referring back to my original five dice-scores which were 10, 8, 6, 7 and 4, these corresponded respectively to the marble finishing up 3 right, 1 right, 1 left, directly under and 3 left of the funnel. So, with the funnel stuck at 30, you can immediately see that the marble finishes up at 33, 31, 29, 30 and 27 respectively:

Stage number	1	2	3	4	5	6
is at	**30**	**30**	**30**	**30**	**30**	**30**
Dice-score =	10	8	6	7	4	
From , then goes	3R	1R	1L	⇓	3L	
Outcome: is at	**33**	**31**	**29**	**30**	**27**	
is still at	**30**	**30**	**30**	**30**	**30**	

Note that, since we're using the same sequence of dice-scores as before, we're also bound to repeat the identical sequence of moves in the "From , then goes" rows (my moves are in green) from page 44 *[WB 38–39]*. So we won't even need to include the rows of dice-scores in the remaining tables! If it isn't convenient for you to make a copy of the table on page 44 *[WB 38–39]*) or detach it from your folder, the sensible way to carry out some manual copying to minimise the amount of page-turning is to copy the sequence of the "From , then goes" moves (3R, 1R, 1L, ⇓, 3L, ...) from your first table onto a separate sheet of paper and then to copy *that* sequence straight into the new table on page 47 *[WB 40]*). Furthermore, also remember to do this first when you tackle the remaining two Rules (on pages 50 and 54 *[WB 42 and 44]* respectively). Alternatively, why not do it right now in all three cases so that you don't have that chore waiting for you later on!

As you can see, with Ford's Second Strategy (Rule 1) you also no longer need the two rows which were under each yellow-shaded row in the Ford's First Strategy (Rule 2) version of the book-keeping (see page 42 and then your table on page 44 *[WB 38–39]*): they were only there to work out where that strategy (Rule 2) would move the funnel. Now, of course, the funnel stays put at 30. So, with everything being so much simpler than before, I think you'll find it very easy to do the book-keeping this time!

This table is also on Workbook page 40.

Stage number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
is at	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**
From , then goes	3R	1R	1L	⇓	3L									
Outcome: is at	**33**	**31**	**29**	**30**	**27**									

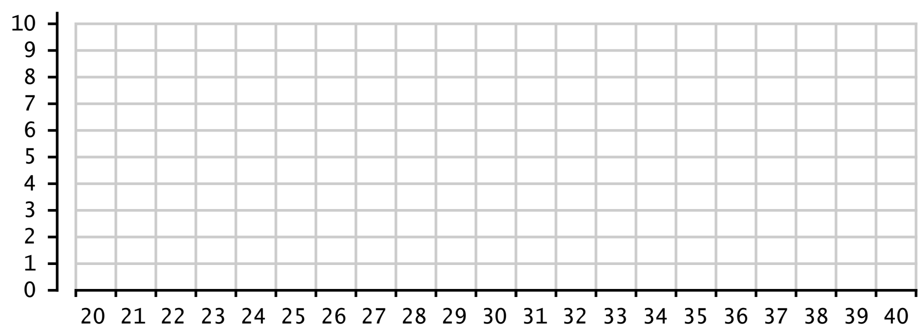
Stage number	15	16	17	18	19	20	21	22	23	24	25	26	27
is at	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**
From , then goes													
Outcome: is at													

Stage number	28	29	30	31	32	33	34	35	36	37	38	39	40
is at	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**	**30**
From , then goes													
Outcome: is at													

So now summarise your new data in a histogram as before.

Tally chart

Histogram



Discussion

Compare the two histograms you've drawn on pages 45 and above *[WB 39 and 41]*. As you'll probably have realised already, the histogram that you've just constructed is somewhat more tightly clustered around the desired value of 30 than the previous one was. The difference may not be dramatic, but it's certainly there. So this process performs *better* than the previous one: there is *less* variation. Yet this process was the "lazy" one. It is much less complicated, much quicker and easier to operate. With this strategy (Rule 1) we haven't "tweaked" the process at all. We've put in less effort—and got better performance! With Ford's original strategy (Rule 2), we worked harder but got worse performance. Some people seem to think that there's a kind of rule of life stating that the harder we work, the better are our results. But, as Dr Deming often pointed out, it's better to do *nothing* than to do the *wrong thing*! He was definitely an advocate of "Work smarter, not harder"! (See the 10th of the 14 Points on Day 5 page 8 *[WB 76]*.)

I asked you a question in Activity 3–c (page 6 *[WB 33]*) about the Ford automatic compensation example. Hopefully, you will now understand why I promised you would be able to answer that question before completing this Major Activity. Unless you produced a very peculiar set of scores from throwing your dice, you should see some similarities between the two histograms you've just drawn and the histograms which came from the Ford example on page 5—not in detail, of course, but broadly in terms of comparing one with the other. As there, the first histogram is more widely spread out than the second one, showing that the harder work (and, in Ford's case, the extra expense of the automatic compensation equipment) actually *increased* the variation—making things worse, as Dr Deming pointed out.

So, to summarise, we have just seen that it is Rule 1 of the Funnel which has produced the better results. It was the original strategy, Rule 2, which produced the poorer results. The statistician's advice to Ford (middle of page 5) was indeed wise.

Let's now move on to ...

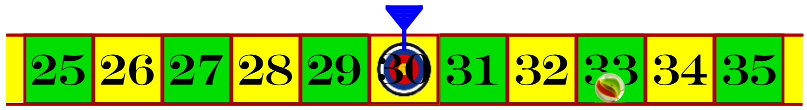
Rules 3 and 4 of the Funnel

The motivation for the various Rules of the Funnel is discussed in *DemDim* Chapter 5, so there is no need for me to go into great detail here. The important issue here is how well they work. So I'll summarise the motivations for Rules 3 and 4 quite briefly. Firstly, Rule 2 was a relatively complicated operation—you might have had to concentrate quite hard to get it right. Rule 3 is a rather easier and (at least at first sight) an apparently rather innocent variant of Rule 2. Rule 4 is quite different from both of them: it concentrates on trying to minimise the average *short-term* variation. The latter is, of course, an interesting mixture of good and bad. It is certainly good to reduce variation, but is it wise to do so only in the short term? Let's carry out similar procedures as previously, but now for Rules 3 and 4, and see what happens.

Recall that in Rule 2 the distance and direction between the target and where the marble comes to rest is described in red just below the yellow-shaded rows (refer back to page 42). Rule 2 then shifts the funnel in the *opposite* direction by that same distance *from its current position*. The difference between Rules 2 and 3 is that, in Rule 3, the funnel is instead placed on the opposite side of *the target* and at that same distance from it. E.g., suppose that the current outcome (position of marble) is 32 (i.e. 2 *above* or to the *right* of target). Then, with Rule 3, the funnel simply gets placed 2 squares to the *left* of target, i.e. at 30 – 2 = 28—*irrespective of where it was before*. Or, if the current outcome was 27 (3 *below* or *left* of target), the funnel is then placed 3 squares to the *right* of target, i.e. at 30 + 3 = 33. This is indeed easier than Rule 2: you'll soon see that the marble and the funnel now simply finish up symmetrically positioned either side of the target—the funnel's *previous position* in no way affects this decision. In case of doubt, just refer to this little table:

If marble is at	…	25	26	27	28	29	30	31	32	33	34	35	…
you move funnel to	…	35	34	33	32	31	30	29	28	27	26	25	…

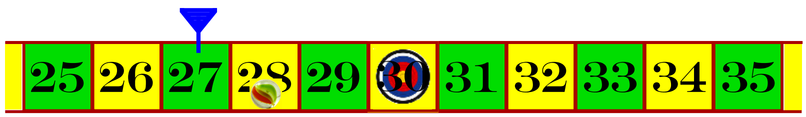
So, using your own track, let's walk through the first few stages with Rule 3. As usual, we'll start with the funnel at 30. With the first dice-score being 10 corresponding to the marble finishing up at 3 right of the funnel, the marble's first finishing position is again 33:



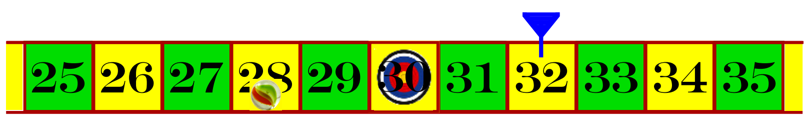
Now (checking if necessary in that little table), with the marble being at 33 (3 *right* of), you move the funnel to the other side of the *target* and at that same distance from it, i.e. to 27 (3 *left* of).

You'll realise that this is actually the same as the initial move with Rule 2; but that is simply because the funnel was initially at the target, i.e. in *both* cases the funnel moved from a starting position of 30.

The second dice-score was 8 corresponding to the marble dropping at 1 right of the funnel, and so the next position of the marble is 28:



Since the marble is now at 28 (2 *left* of), Rule 3 moves the funnel to 32 (2 *right* of), i.e. to the opposite side of the target and at that same distance from it. Remember, at this and every stage, you simply move the funnel so that it and the marble are symmetrically placed either side of the target :



The third dice-score was 6 corresponding to the marble dropping at 1 left of the funnel. So the marble now finishes up at 31. 31 is 1 *right* of , and thus you move the funnel to 29 (1 *left* of), giving:

And so on. The fourth and fifth dice-scores were 7 and 4 corresponding respectively to the marble falling directly under the funnel and 3 to the left of the funnel. Carefully confirm with your track that the next finishing positions are: marble under funnel at 29, funnel moved to 31, marble at 28, and funnel moved to 32.

Just for interest, take a moment's thought to guess what will happen in the long term as you take Rule 3 through its 40 stages. Then, later on, see if you've guessed correctly!

Now, as before, continue your book-keeping using your own dice-scores. Remember that if finishes exactly on-target at 30 then you indicate this by a √ in the "As relative to is …" row. However, with Rule 3, the consequence is that you then also move to 30, i.e. directly above the marble; I suggest it would therefore be logical for you to indicate this by ⇑ in the "… place relative to " row.

If you haven't already done so, start by copying your "From , then goes" rows from page 44 *[WB 38–39]* through this whole table. Next, check through the first five stages again using your track, and then carry on with the rest of your stages.

Pages 50–51 are also on Workbook pages 42–43.

Stage number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
is at	**30**	**27**	**32**	**29**	**31**	**32**								
From , then goes	3R	1R	1L	⇓	3L									
Outcome: is at	**33**	**28**	**31**	**29**	**28**									
As relative to is ...	3R	2L	1R	1L	2L									
... place relative to	3L	2R	1L	1R	2R									
So is now at	**27**	**32**	**29**	**31**	**32**									

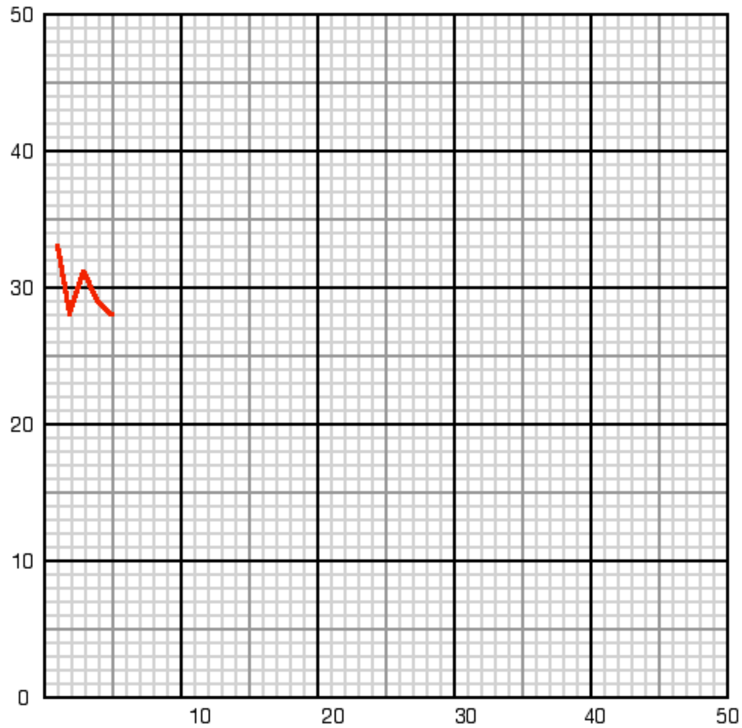
NB By now you may already have realised that you can again short-circuit the two rows beneath the yellow-shaded **"Outcome"** row, in this case by using the symmetric property emphasised on the previous two pages. This is equivalent to immediately writing in the bottom "So is now at" row the number which, when added to the **"Outcome"** number, totals 60. That's quicker! Confirm that it's true in the first part of the table above, and then continue to use it if you like it. As soon as you are happy with it, you can skip the two rows under the yellow-shaded row and proceed straight down to the bottom row. Or if you are not happy, carry on using those two rows.

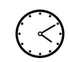
Stage number	15	16	17	18	19	20	21	22	23	24	25	26	27
is at													
From , then goes													
Outcome: is at													
As relative to is ...													
... place relative to													
So is now at													

Stage number	28	29	30	31	32	33	34	35	36	37	38	39	40
is at													
From , then goes													
Outcome: is at													
As relative to is ...													
... place relative to													
So is now at													

And now draw the run chart, continuing on from my first five points. Again, as with the histograms, there's no need to be too artistic about it—please yourself whether or not you use a ruler!

Rule 3





Recall that the very first stage in Rule 3 was identical to that in Rule 2 but, as was pointed out at the time, this was simply because the funnel was initially at its "sensible" value of 30. I imagine, however, that you subsequently discovered things changed with Rule 3 … . You may recall Peter Worthington's observation (page 29) that a zig-zag pattern is *not* random variation!

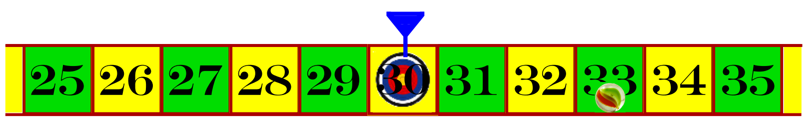
(Move on to the description of Rule 4 on the next page.)

Lacking basic understanding of variation, management may with best intentions have used Rules 2 and/or 3 in the hope that they would produce better results than Rule 1. They have failed dismally: instead they have considerably *increased* variation. Therefore they now try something completely different. Since it is good to reduce variation, they strive to minimise variation *at least in the short term*. Certainly there are circumstances where this may seem quite reasonable in practice. For example, if you don't get exactly what you want from a supplier, but at least what you receive doesn't change much from one delivery to the next, it is quite possible that you can learn to "cope" (for a while) with whatever arrives. The situation that really throws you is when you get dramatic changes from one delivery to the next (which presumably is what happened to you sooner or later during Rule 3)—that's part of the motivation for trying Rule 4.

So, what do we do with the funnel in Rule 4? Suppose the current outcome is 33. The thinking is now that, for *low short-term variation*, we would like the next outcome to be 33 again or pretty close to it. How do we achieve that? Easy! Place the funnel exactly where the marble has just finished up, i.e. at 33. You will have noticed that (again unless you've had a very unusual sequence of dice-scores) most of the time the marble was finishing up quite close to the funnel—say, no more than 2 squares away from it. So, in this case, the chances are that your next outcome will now be somewhere between 31 and 35—nicely consistent with the aim of achieving low short-term variation, and certainly generally more appealing than what Rule 3 eventually produced!

Take a look at my fantastic outcomes (the yellow-shaded row) near the top of page 54 *[WB 44]*! Hardly any variation for a while; and then, when there is a larger move, we find ourselves right back on the target of 30! (As you might realise, that was a very happy accident.) I wonder what will happen in the longer term ...

So, just to be sure, let's again walk through the first few stages. As always, with the funnel initially being placed at 30 and the first dice-score being 10, the first outcome is again 33:



So, as argued near the bottom of page 50, Rule 4 then places the funnel directly above the marble at 33:

The second dice-score was 8 corresponding to the marble dropping at 1 right of the funnel, and so the next position of the marble is 34:

and then the funnel immediately follows it there:

The third dice-score was 6 which corresponds to the marble dropping at 1 left of the funnel, i.e. back to 33, and so Rule 4 moves the funnel back to 33 as well:

And so it continues. The fourth and fifth dice-scores were 7 and 4 corresponding respectively to the marble falling directly under the funnel and then 3 to the left of the funnel. Confirm that the next positions are Stage 4: both marble and funnel still at 33, and Stage 5: both marble and funnel finish up at 30. So the picture stays unchanged at the fourth stage and, at the fifth stage, both the marble and then the funnel end up at the target of 30—the "happy accident" to which I referred on page 52.

So off you go for the final time. Again if you didn't do so when I first suggested it, you'll need to start by copying over all your "From , then goes" rows one final time. Then, after you've generated all your Rule 4 **outcome**s, finish off by sketching your run chart. As usual, I've started the run chart for you on the graph-paper. Also, as with Rule 3, do have an early guess at what you think will happen in the long term!

Rule 4 is one of the easier strategies to carry out, so you can stop using your track as soon as you realise you don't need it any more. That may be very soon!

(Move on to the table on page 54 [WB 44].)

Pages 54–55 are also on Workbook pages 44–45.

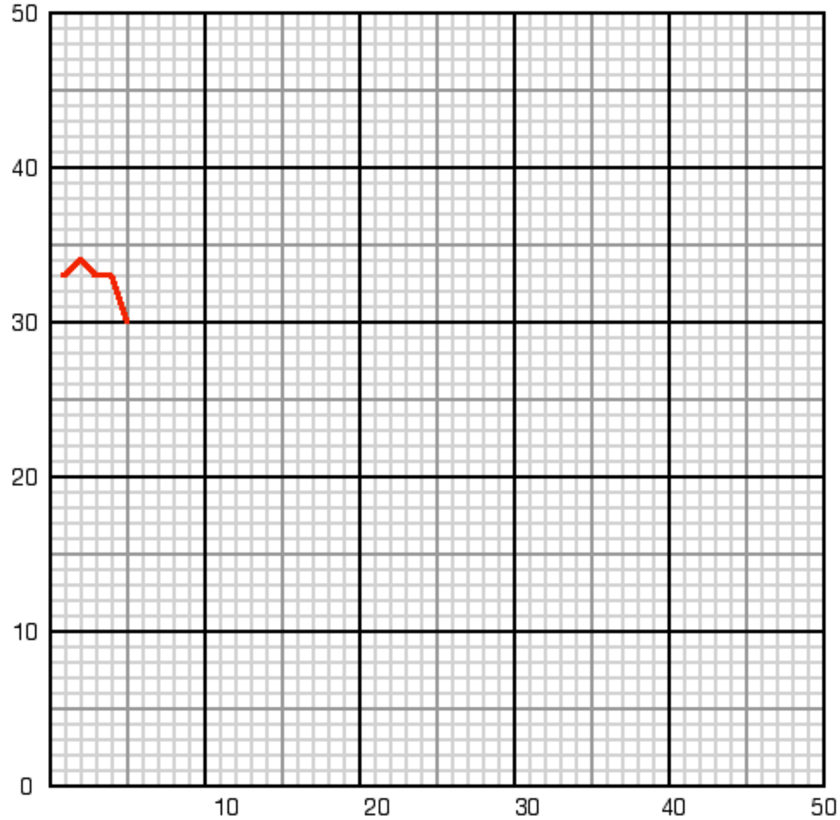
Stage number	1	2	3	4	5	6	7	8	9	10	11	12	13	14
is at	**30**	**33**	**34**	**33**	**33**	**30**								
From , then goes	3R	1R	1L	⇓	3L									
Outcome: is at	**33**	**34**	**33**	**33**	**30**									
So move to	**33**	**34**	**33**	**33**	**30**									

Stage number	15	16	17	18	19	20	21	22	23	24	25	26	27
is at													
From , then goes													
Outcome: is at													
So move to													

Stage number	28	29	30	31	32	33	34	35	36	37	38	39	40
is at													
From , then goes													
Outcome: is at													
So move to													

(And now move straight on to the run chart on the next page:)

Rule 4



Summary

As has been obvious for a while (if not from the very start!), you know that, assuming we are "stuck" with the amount of variation represented by throwing the two dice, the best thing to do is to put the funnel at the desired value of 30 and *leave it there*. Moving the funnel anywhere at any time following any "strategy" will "only make things worse". I trust that sounds familiar!

But putting the funnel at 30 and leaving it there, i.e. Rule 1, is *not* what many managers (amongst others) are keen on doing. A more familiar scenario is that each result is compared with the target and some apparently appropriate action is taken according to the difference. Specifically, if a result is above the target then an action is taken to try to lower the next value to try to "get it right"; or, if the result is below the target, an action is taken to try to raise the next value. As Dr Deming described it—see page 2—this was indeed "a noble aim". But recall he then also pointed out that "There was only one little trouble". Doubtless you can recall what that "little trouble" was! This is precisely what Rule 2 did. As pointed out earlier, a very attractive argument for Rule 2 is that, *if only* the funnel had already been where Rule 2 now puts it, we would have just hit the target spot on. That's a supreme example of "being wise after the event". Have you heard of JIT (Just In Time)? This is JTL (Just Too Late)!

Since Rule 2 was seen not to do the trick, Rule 3 was tried: Rule 3 is an alternative and simpler version of the same idea. It turned out to be disastrous! Do people then conclude that it would have been better not to do anything at all but to have stuck with Rule 1? Somehow it doesn't seem to be very politically acceptable in management meetings or committee meetings and the like to say that all of our ideas so far have been wrong and it would have been better if we had just done nothing. So a completely different kind of attempt is then tried: Rule 4. For a while it may look as if things are going well: as we have seen, it concentrates on reducing short-term variation as much as possible. It succeeds in doing that. The consequence is that it may actually look better than Rule 1 in the short term! But not for long … . Actually you may have been suspicious that trouble might lie ahead from the simple fact that (except for initially placing the funnel at 30) the operation of Rule 4 totally ignores the target!

What do Rules 2, 3 and 4 do? They show how, in Dr Deming's terms, people (especially management) spend so much of their time *tampering with* the system (= *thinking* they may be doing something useful but, in fact, making performance worse) rather than *improving* the system (= making performance better). I think that, having carried out this Major Activity, that word "tampering" may also mean more to you now than it used to!

Some may find that phrase "it would have been better if we had just done nothing" to be somewhat startling! So, to be sure of its context, let's recall a sentence from page 48: "But, as Dr Deming often pointed out, it's better to do *nothing* than to do the *wrong thing*!". That's the point. Obviously it's *not* better to do *nothing* rather than doing the *right thing*! But the vital message to strike home is that, when you have a stable process, the *only* "right thing" to do is to investigate that process carefully in order to discover what is causing some of its current variation and then to make some changes to the process which will reduce that variation (and not only in the short term, for that's Rule 4). *There is no other way.*

Finally in this Major Activity, please take a brief look at a couple of runs of the experiment that I have carried out using my computer simulation program. There is also some more discussion there on the Funnel Experiment in general and on this Major Activity in particular. See Appendix pages 15–18.

In Part A of the Optional Extras section we shall examine in detail what happens to control charts when presented with the data from the two simulations of the Funnel Experiment studied in the Appendix. However, particularly in case you decide that that optional material is not for you, Activities 3–i and 3–j consider some general effects of putting data from the four Rules of the Funnel onto control charts.

Activity 3–i is also on Workbook page 46.

ACTIVITY 3–i

Having spent much of this morning on beginning to get used to control charts and now this afternoon on the Funnel Experiment, this is a useful Activity which involves both of them. However, if you are a Stats-level 0 student then this possibly isn't for you.

Let's first précis the "How Do We Compute Those Control Limits—and Why?" section on pages 13–15 as follows:

> "Control limits need to indicate the range over which the data will vary when the process is in statistical control: so that, if and when data go outside those limits, we have evidence that the process may well be out of statistical control. But suppose the process is *out of* statistical control when we collect those data. The method we use is based on 'moving ranges'. Obviously, if many of these moving ranges are large then high variation is indicated; whereas if the moving ranges are mostly small then low variation is indicated. Using moving ranges works pretty well in mitigating the contamination effects of many kinds of special causes. There are a few exceptions. Two important exceptions that one needs to be able to recognise are illustrated with data generated in the Funnel Experiment, and so we shall see those this afternoon."

With these thoughts in mind, imagine that you are computing control limits (using moving ranges) from data that are being generated from each of the four Rules of the Funnel in turn. How do you think those data will affect the control limits, and what would happen if you extended those limits into the future during which the same Rule is in operation?

Rule 1

Rule 2

Rule 3

Rule 4

(For discussion see Appendix pages 18–20.)

DemDim Chapter 5 describes the original version of the Funnel Experiment and shows some typical results from it. The chapter is quite short: less than 12 pages, and you should now find it to be a very quick and easy read. So, for completeness, **do now please read it**—you will soon see the analogies with what you have produced during your experiment. In the second half of the chapter you will find some real-life illustrations of the messages coming from the Funnel Experiment—they are but a few of many.

Activity 3–j is also on Workbook page 47.

ACTIVITY 3–j

Now that you have completed the Major Activity and after reading *DemDim* Chapter 5, it would be good if you could spend the final few minutes summarising some illustrations of your own. You will find further suggestions in the relevant discussion in the Appendix—but don't look at it just yet!

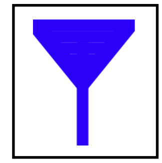
> It is often not possible to differentiate between Rules 2 and 3 when suggesting examples. As we have seen, in effect the difference between them depends on whether the tampering is done comparatively sensibly or completely stupidly! Also, the performances of some practical illustrations are worse than Rule 2 but not as berserk as Rule 3! I would recommend therefore that, in addition to a list of possibilities for Rule 4, you simply compile one other list to cover both Rules 2 and 3 to thus include any kind of "zig-zag" or "swinging the pendulum" compensation effect. Further, the Rule 4 list does not need to be restricted to a *strict* version of Rule 4. For example, a photocopy of a photocopy of a photocopy of … is similar to Rule 4 except that it continually moves away from the target (the original copy) rather than being able to temporarily move back toward it.

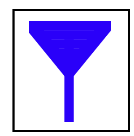
(For some final examples see Appendix pages 20–21.)

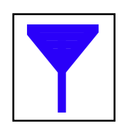
As mentioned on page 56, the data from the two runs of the Funnel Experiment that are examined in the Appendix are studied using control charts in Part A of the Optional Extras section. However, we are now at the end of Day 3 and therefore to look at that optional extra material will indeed have to be an "out-of-hours" activity!

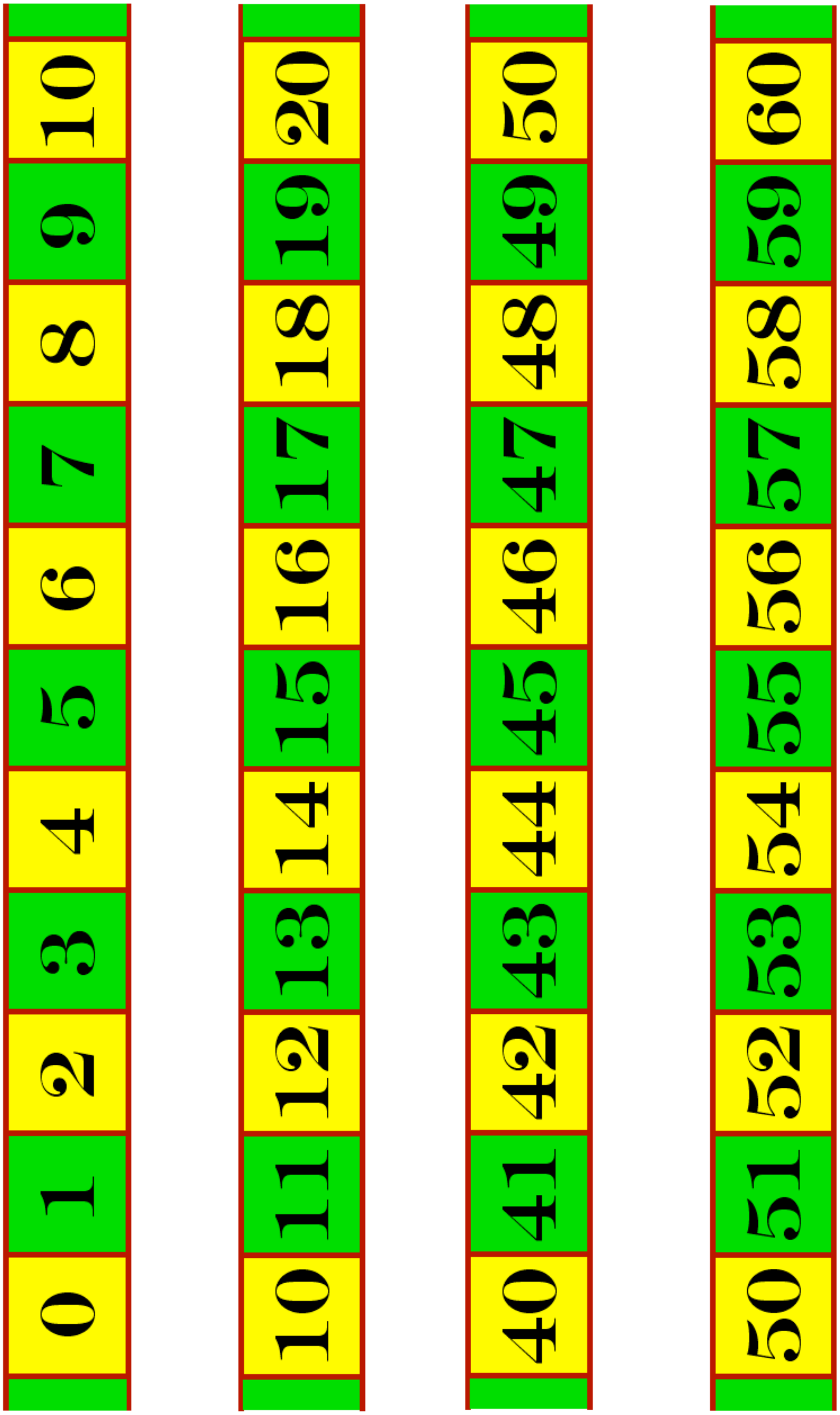

30
29
28
27
26
25
24
23
22
21
20


40
39
38
37
36
35
34
33
32
31
30








0 1 2 3 4 5 6 7 8 9 10
10 11 12 13 14 15 16 17 18 19 20
40 41 42 43 44 45 46 47 48 49 50
50 51 52 53 54 55 56 57 58 59 60

Approvals, Acknowledgments and Information

[a] (page 5) The Ford Motor Company example and diagrams are included with the approval of Bill Scherkenbach.

[b] (page 32) This and all other quotations from *Understanding Statistical Process Control* have been reproduced with the approval of SPC Press Inc.