Myths About Process Behavior Charts

How to avoid some common obstacles to good practice

Donald J. Wheeler

The simplicity of the process behavior chart can be deceptive. This is because the simplicity of the charts is based on a completely different concept of data analysis than that which is used for the analysis of experimental data. When someone does not understand the conceptual basis for process behavior charts they are likely to view the simplicity of the charts as something that needs to be fixed. Out of these urges to fix the charts all kinds of myths have sprung up resulting in various levels of complexity and obstacles to the use of one of the most powerful analysis techniques ever invented. The purpose of this paper is to help you avoid this complexity.

Myth One: It has been said that the data must be normally distributed before they can be placed on a process behavior chart.

In discussing this myth some historical background may be helpful. Walter Shewhart published his "*Economic Control of Quality of Manufactured Product*" in 1931. When the British statistician E. S. Pearson read Shewhart's book he immediately felt that there were gaps in Shewhart's approach, and so he set out to fill in these perceived gaps. The result was Pearson's 1935 book entitled "*The Application of Statistical Methods to Industrial Standardization and Quality Control.*" In this book Pearson wrote on page 34: "Statistical methods and tables are available to test whether the assumption is justified that the variation in a certain measured characteristic may be represented by the Normal curve."

After reading Pearson's book, Shewhart gave a series of lectures that W. Edwards Deming edited into Shewhart's 1939 book, "*Statistical Method from the Viewpoint of Quality Control.*" In choosing this title Shewhart effectively reversed Pearson's title to emphasize that his approach solved a real problem rather than being a collection of techniques looking for an application. On page 54 of this second book Shewhart wrote: "*we are not concerned with the functional form of the universe, but merely with the assumption that a universe exists.* [Italics in the original]." Here Shewhart went to the heart of the matter. While Pearson essentially assumed that the use of a probability model would always be justified, Shewhart created a technique to examine this assumption. The question addressed by a process behavior chart is more basic than "What is the shape of the histogram?" or "What is the probability model?" It has to do with whether we can meaningfully use any probability model with our data.

Shewhart then went on to note that having a symmetric, bell-shaped histogram is neither a prerequisite for the use of a process behavior chart, nor is it a consequence of having a predictable process. Figure 1 shows Shewhart's Figure 9 from the 1931 book. He characterized these data as "at least approximately [in] a state of control." This skewed histogram is certainly not one that anyone would claim to be "normally distributed." So, while Shewhart had thoroughly examined this topic in his 1931 book, his approach was so different from traditional statistical thinking that Pearson and countless others (including this author on his first reading) completely missed this crucial point.

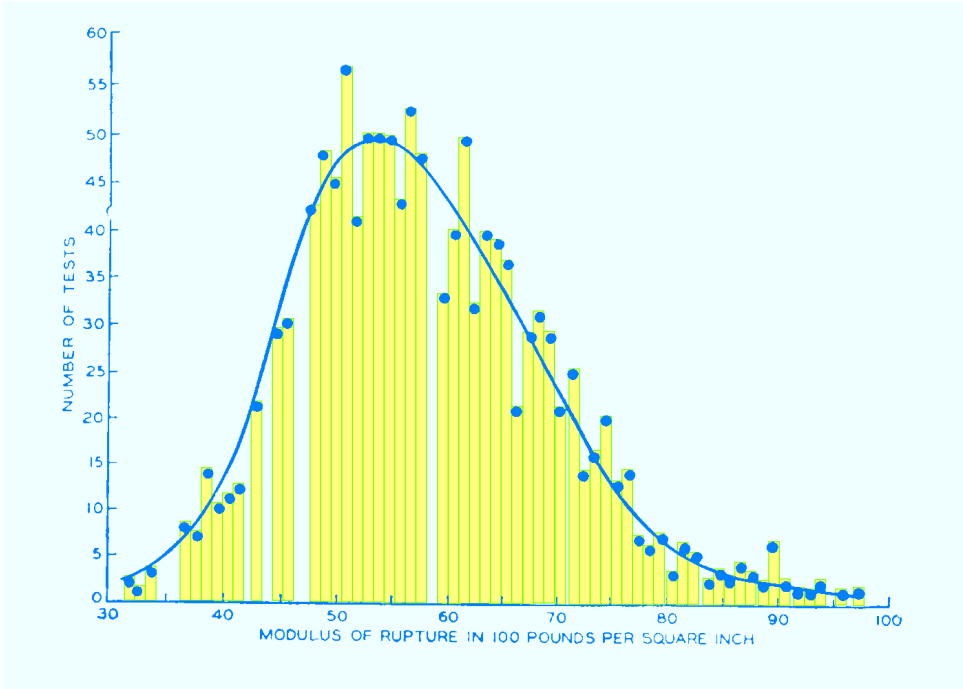

Figure 1: Shewhart's Figure 9: Variability in Modulus of Rupture of Clear Specimens of Green Sitka Spruce Typical of the Statistical Nature of Physical Properties

To begin to understand how a process behavior chart can be used with all sorts of data we need to begin with a simple equation from page 275 of Shewhart's 1931 book:

$$\int_A^B f(x)\ dx \quad = P$$

Shewhart described two completely different approaches this equation. The first of these approaches I call the statistical approach since it describes how we approach statistical inference: (1) Choose an appropriate probability model *f(x)* to use; (2) Choose some small risk of a false alarm ($1 - P$) to use; (3) Find the exact critical values A and B for the selected model that correspond to this risk of a false alarm; (4) Then use these critical values in your analysis. While this approach makes sense when working with *functions* of the data (i.e. statistics) for which we know the appropriate probability model, it encounters a huge problem when it is applied to the original data. As Shewhart pointed out, we will *never* have enough data to uniquely identify a specific probability model for the original data. In the mathematical sense all probability models are limiting functions for infinite sequences of random variables. This means that they can never be said to apply to any finite portion of that sequence. This is why any assumption of a probability model for the original data is just that—an assumption that cannot be verified in practice. (While lack-of-fit tests will sometimes allow us to falsify this assumption, they can never verify an assumed probability model.)

So what are we to do when we try to analyze data? Shewhart suggested a different approach for the analysis of original data. Shewhart's approach to the equation above was: (1) Choose some generic critical values A and B for which (2) the risk of a false alarm ($1 - P$) *will be*

reasonably small (3) *regardless* of what probability model *f(x)* we might choose, and (4) use these generic critical values in your analysis. This approach changes what is fixed and what is allowed to vary. With the statistical approach the alpha-level is fixed, and the critical values vary to match the specific probability model. With Shewhart's approach it is the critical values that are fixed (three-sigma limits) and the alpha-level that is allowed to vary. This complete reversal of the statistical approach is what makes Shewhart's approach so hard for those with statistical training to understand.

Once you see the difference in these two approaches you can begin to see why Pearson and others have been concerned with the probability model *f(x)*, why they have sought to maintain a fixed alpha level (1 – *P*), and why they have been obsessed with the computation of exact values for *A* and *B*. And more recently, you can see how others have become obsessed with transforming the data prior to placing them on a process behavior chart. Their presuppositions prevent them from understanding how Shewhart's choice of three-sigma limits is completely independent of the choice of a probability model. In fact, to their way of thinking, you cannot even get started without a probability model. Hence, as people keep misunderstanding the basis for process behavior charts, they continue to recreate Myth One.

Myth Two: It has been said that process behavior charts work because of the central limit theorem.

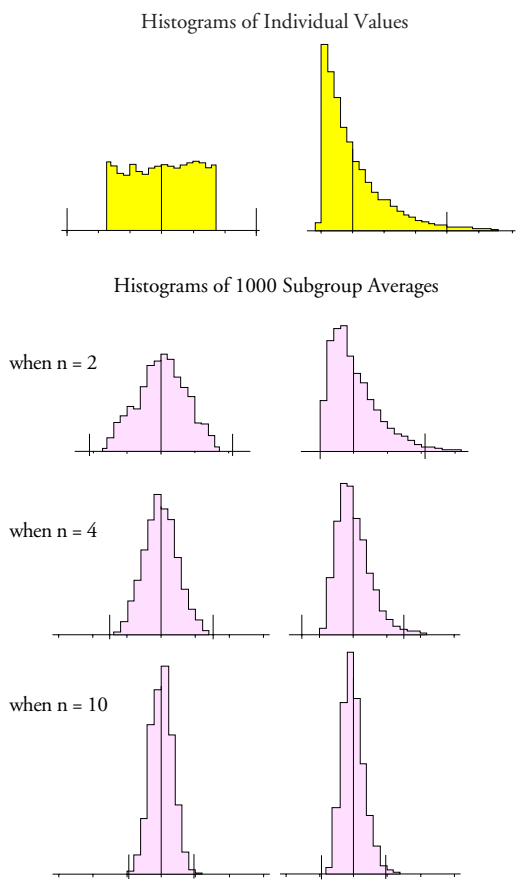

Figure 2: The Central Limit Theorem for Subgroup Averages

The central limit theorem was published by Laplace in 1810. This fundamental theorem shows how, regardless of the shape of the histogram of the original data, the histograms of subgroup averages will tend to have a "normal" shape as the subgroup size gets larger. This is illustrated in Figure 2 where the histograms for 1000 subgroup averages are shown for each of three different subgroup sizes for data obtained from two completely different sets of original data. There we see that even though the histograms for the individual values differ, the histograms for the subgroup averages tend to look more alike and become more bell-shaped as the subgroup size increases.

Many statistical techniques that are based on averages utilize the central limit theorem. While we may not know what the histogram for the original data looks like, we can be reasonably sure that the histogram of the subgroup averages may be approximated by a normal distribution. From this point we can then use the statistical approach outlined in the preceding section to carry out our analysis using the subgroup averages.

However, while we have a central limit theorem for subgroup averages, there is no central limit theorem for subgroup ranges. This is illustrated in Figure 3 where we see the histograms of the subgroup ranges obtained from two different sets of original data. Each histogram shows the ranges of 1000 subgroups, for each of three subgroup sizes, obtained from each of the two data sets shown. As the subgroup size increases the histograms for the subgroup ranges become more dissimilar and do not even begin to look bell-shaped.

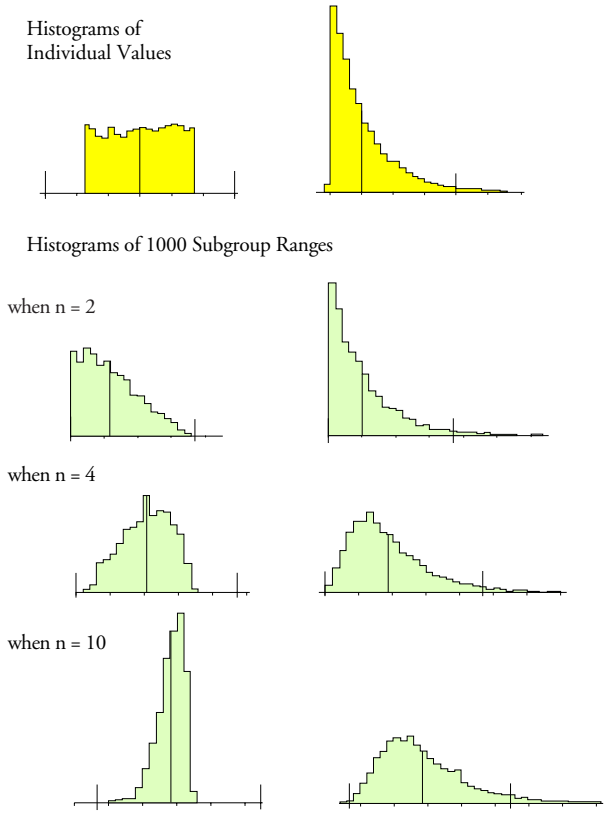

Figure 3: The Lack of a Central Limit Theorem for Subgroup Ranges

Therefore, Myth Two has no basis in reality. If the central limit theorem was the foundation for process behavior charts, then the range chart would not work.

Rather, as we saw in the preceding section, Shewhart chose three-sigma limits to use with the process behavior chart simply because, when the data are homogeneous, these limits will bracket virtually all of the histogram regardless of the shape of that histogram. Three-sigma limits are shown on each of the 16 histograms in Figures 2 and 3. There they bracket better than 98 percent of each histogram, leaving less than a 2 percent chance of a false alarm in each case. In practice, as long as $(1-P)$ is known to be small, you do not need to know the exact risk of a false alarm. This means that when you find a point outside the limits of a process behavior chart the odds are very good that the underlying process has changed and you will be justified in taking action. Three-sigma limits provide you with a suitably conservative analysis without requiring a lot of preliminary work. It is this conservative nature of three-sigma limits that eliminates the need to appeal to the central limit theorem to justify the process behavior chart.

Undoubtedly, Myth Two has been one of the greatest barriers to the use of process behavior charts with management data and process-industry data. Whenever data are obtained one-value-per-time-period it will be logical to use subgroups of size one. However, if you believe Myth Two you will feel compelled to average something in order to invoke the blessing of the central limit theorem, and the rationality of your data analysis will be sacrificed to superstition. The conservative nature of three-sigma limits allows you to use the chart for individual values with all sorts of original data without reference to the shape of the histogram.

Myth Three: It has been said that the observations must be independent—data with autocorrelation are inappropriate for process behavior charts.

Again we have an artificial barrier to the use of a process behavior chart which ignores both the nature of real data and the robustness of the process behavior chart technique. Virtually all data coming from a production process will display some amount of autocorrelation. Autocorrelation is simply a measure of the correlation between a time series and itself. A positive autocorrelation (lag one) simply means that the data display two characteristics: (1) successive values are generally quite similar while (2) values that are far apart can be quite dissimilar. These two properties mean that when the data have a large positive autocorrelation the underlying process will be changing. To illustrate this property I will use the data from Table 2, page 20 of Shewhart's 1931 book. These data are the measured resistances of insulation material. These data have an autocorrelation of 0.549, which is detectably different from zero (also known as significantly different from zero). While Shewhart organized these data into 51 subgroups of size four and placed them on an average chart, it could be argued that this subgrouping obscures the effects of the autocorrelation upon the chart. To avoid this problem I have placed these 204 data on an *XmR* chart in Figure 4.

Shewhart found 8 averages outside his limits. We find 14 individual values and 7 moving ranges outside our limits. So both Shewhart's average chart and our *XmR* chart tell the same story. This process was not being operated predictably.

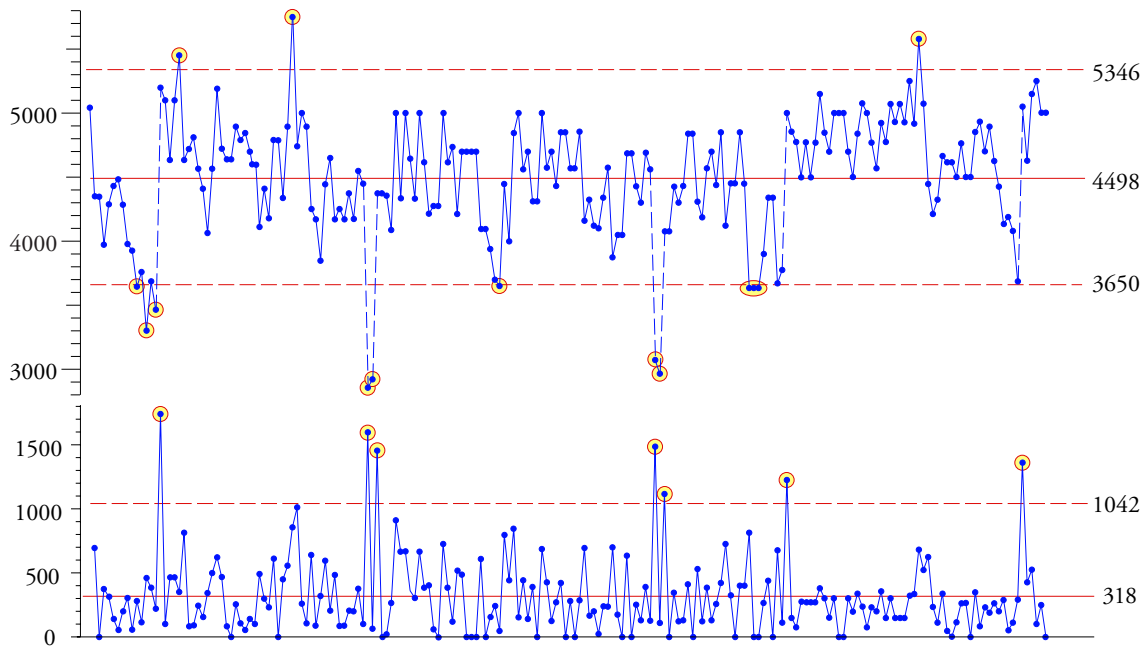

Figure 4: *XmR* Chart for 204 Resistances from Shewhart (1931) Page 20

As they found the assignable causes and took steps to remove their effects from this process they collected some new data. These data, shown in Figure 5, show no evidence of unpredictable behavior. Notice that the new limits are only 60% as wide as the original limits. By removing the assignable causes of exceptional variation they not only got rid of the process upsets and the extreme values, but they also removed a substantial amount of process variation. The autocorrelation for the data in Figure 5 is 0.091, which is not detectably different from zero.

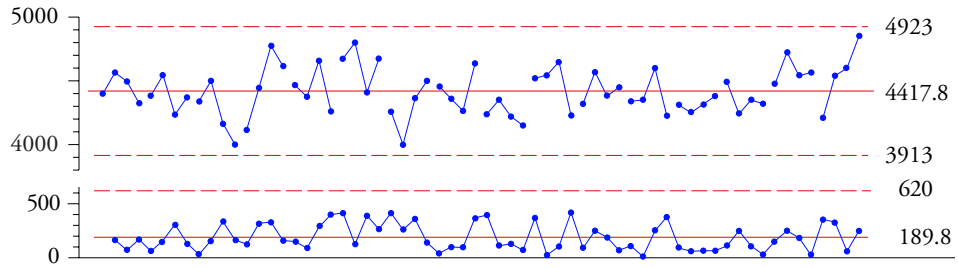

Figure 5: *XmR* Chart for 64 Additional Resistances from Shewhart (1931) Page 20

This example illustrates an important point. Whenever the data have a substantial autocorrelation the underlying process will be changing, and vice-versa, when the process is moving around the data will tend to have an autocorrelation that is detectably different from zero. Thus, autocorrelation is simply one way that the data have of revealing that the underlying process is changing. On the other hand, when the process is operated predictably, the data are unlikely to possess a substantial autocorrelation.

Remember that the purpose of analysis is insight rather than numbers. The process behavior

chart is not concerned with creating a model for the data, or whether the data fit a specific model, but rather with using data for making decisions in the real world. To insist that the data be independent is to add something to Shewhart's work that Shewhart was careful to avoid. This example from Shewhart's first book illustrates that process behavior charts have worked with autocorrelated data from the very beginning. Do not let those who do not understand this point keep you from placing your data on a chart because the values might not be independent.

While a complete treatment of the effects of autocorrelation is beyond the scope of this article, the following observation is in order. While it is true that when the autocorrelation gets close to +1.00 or –1.00 the autocorrelation can have an impact upon the computation of the limits, such autocorrelations will also simultaneously create running records that are easy to interpret at face value. This increased interpretability of the running record will usually provide the insight needed for process improvement and further computations become unnecessary.

Myth Four: It has been said that the process must be operating in control before you can place the data on a process behavior chart.

I first encountered this myth when I was refereeing a paper written by a professor of statistics at a land-grant university in the South, which goes to prove my point that even an extensive knowledge of statistics does not guarantee that you will understand Shewhart.

I suspect that the origin of Myth Four is a failure to appreciate that there are correct and incorrect ways of computing the limits for a process behavior chart. (See my January and February Columns from 2010 for more on this topic.) The most common of the incorrect ways of computing limits consists of using three-standard-deviation limits rather than three-sigma limits. While this approach was identified as incorrect on page 302 of Shewhart's 1931 book, it is found in virtually every piece of software available today. While three-standard-deviation limits will mimic three-sigma limits whenever the process is operated predictably, they will be severely inflated when the process is being operated unpredictably. Thus, when someone is using the incorrect way of computing the limits, they might come to believe Myth Four.

Of course, as soon as you believe Myth Four you will begin to look for a way to remedy this perceived defect in the technique. Among the absurdities which have been perpetrated in the name of Myth Four are the censoring of the data prior to placing them on the chart (removing the outliers) and the use of two-standard-deviation limits. (As Henry Neave observed in a letter to the Royal Statistical Society, calculating the limits incorrectly and then using the wrong multiplier is an example of how two wrongs still do not make one right.) Needless to say that these, and all other associated manipulations are unnecessary. The express purpose of the process behavior chart is to detect when a process is changing, and to do this we have to be able to get good limits from bad data.

One of my students had just completed the class and was looking at the archival data they had for their cooling water system. He organized these data into daily subgroups of size five and plotted the averages and ranges for the past 24 days to get the graph shown in Figure 6.

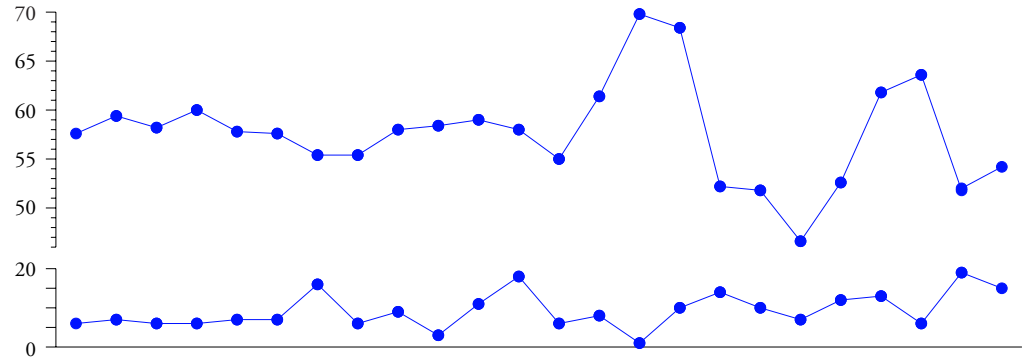

Figure 6: Subgroup Averages and Subgroup Ranges for Cooling Water Pressures

Based on what he saw in Figure 6, Terry decided to use the first half of the data to compute the limits for this chart. When he did this he got the limits shown in Figure 7.

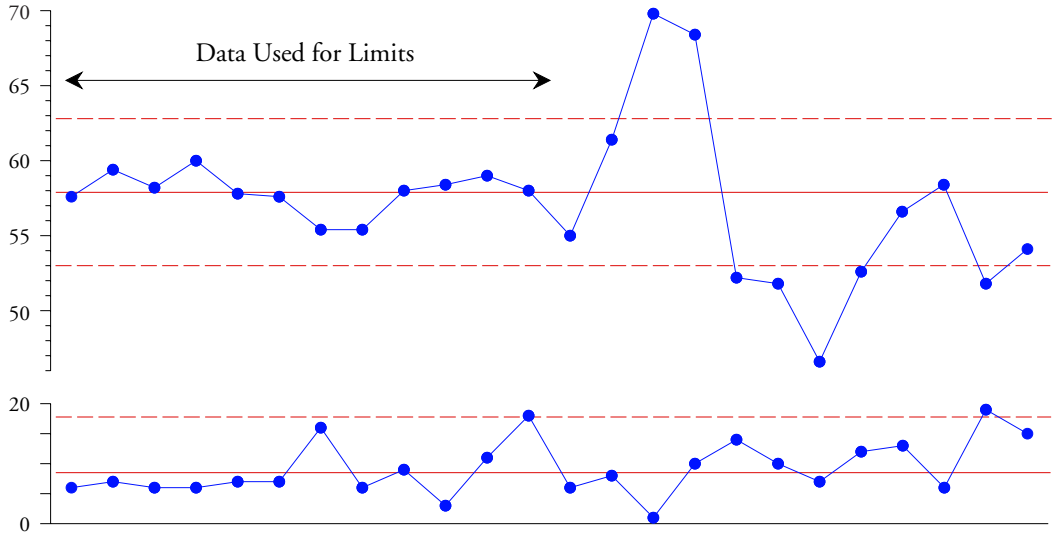

Figure 7: Terry's Average and Range Chart for Cooling Water Pressures

The two high points coincided with the short week prior to the Christmas break, and the drop to the lower points coincided with the January start-up. Based on this chart Terry was able to explain why they were spending over $2 million a year on scrap product. With some minor changes they immediately cut the scrap rate by 70 percent, and by the end of the following year they had cut the scrap rate to 10 percent of what it had been by simply operating their processes more consistently.

But why did Terry only use the first 12 days in computing the data? Because he was afraid that the limits would "blow-up" if he used all of the data. Figure 8 shows the chart of Figure 7 with two sets of limits. The limits shown as solid blue lines were computed using the data from all 24 days. The red dashed lines show Terry's limits.

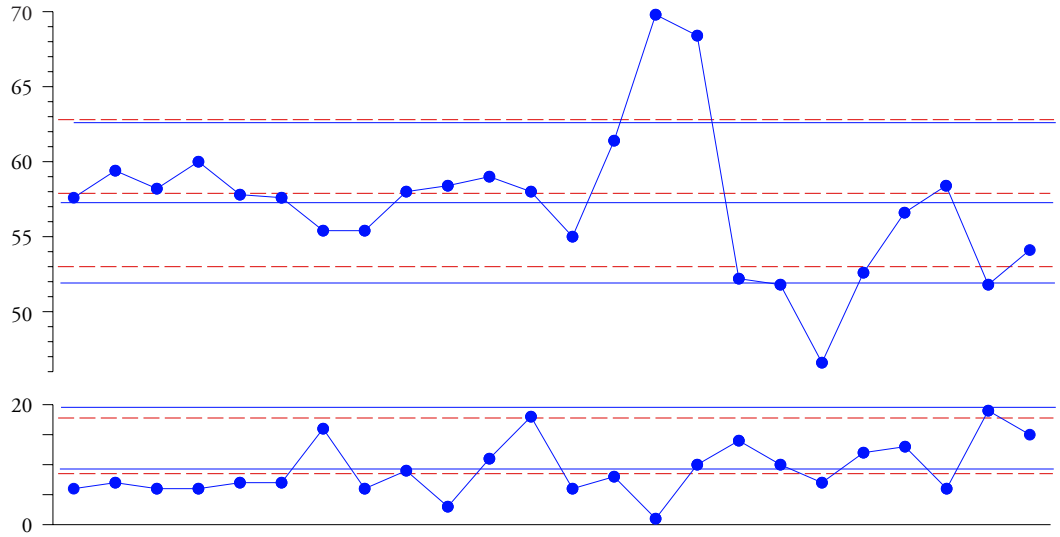

Figure 8: Another Average and Range Chart for Cooling Water Pressures

While the limits do change slightly, the story told by the chart remains the same regardless of which set of limits you use. The purpose of a process behavior chart is to tell the story contained within the data and the limits are simply a means to this end.

Thus, as illustrated by Figure 8, we can compute good limits using bad data. We do not have to wait until the process is "well-behaved" before we compute our limits. The correct computations are robust. And this is why Myth Four is patent nonsense.

SUMMARY

Shewhart's approach to the analysis of data is profoundly different from the statistical approach. This is why people end up with such confusion when they try to "update" Shewhart by attaching bits and pieces from the statistical approach to what Shewhart has already done. As I showed in my March column, "Three Questions for Success," Shewhart provided us with an operational definition of how to get the most out of any process. Nothing extra is needed to make process behavior charts work. We do not need to check for normality or transform the data to make them "more normal." We do not have to use subgrouped data in order to receive the blessing of the central limit theorem before the chart will work. We do not need to examine our data for autocorrelation. And we do not need to wait until our process is "well-behaved" before computing limits. All such "extras" will just mess you up, get in your way, leave you confused, and keep you from using one of the most powerful data analysis techniques ever invented.