Shewhart, Deming, and Six Sigma

Donald J. Wheeler

Since neither time nor space will allow me to cover all that the title above encompasses, I have chosen to focus on selected aspects of the work of Shewhart and Deming and to discuss how these compare with a common element of various six-sigma programs. To this end I will begin with a look at the concept of an operational definition, then turn to what it takes for improvement. This will lead to a distinction between observational studies and experimental studies. Finally I will look at the basic assumptions of virtually all six-sigma programs in the light of the earlier material.

An Operational Definition

In the pre-publication drafts of *Quality, Productivity, and Competitive Position* Dr. Deming wrote:

> "An operational definition consists of (1) a criterion to be applied to an object or a group of objects, (2) a test of compliance for the object or group, and (3) a decision rule for interpreting the test results as to whether the object or group is, or is not, in compliance."

This definition closely parallels Dr. Shewhart's opening statement for his (1939) book *Statistical Method from the Viewpoint of Quality Control*:

> "Broadly speaking there are three steps in a quality control process: the *specification* of what is wanted, the *production* of things to satisfy the specification, and the *inspection* of the things produced to see if they satisfy the specification."

This idea of an operational definition, which Shewhart and Deming popularized from the work of the philosopher C. I. Lewis, provided the seed for what grew into the Shewhart or PDSA Cycle. While the Plan-Do-Study-Act Cycle does form a powerful framework for any improvement effort, it has often been reduced to a checklist to be followed mechanically. This has led to a proliferation of "expanded" PDSA cycles where each of the steps on the checklist are specified in ever increasing detail. But before we go down this path, I would like to back up and generalize a bit.

In Dr. Deming's own conversations, when individuals would start telling him about what they or their organization were planning to do, he would invariably have one of two responses for them: "By what method?" or "How will you know?" Either one of these questions would generally end the conversation since the individual would have no answer. After discerning this pattern to Dr. Deming's responses, it finally occurred to me that these two questions corresponded to the last two parts of an operational definition. This realization, in turn, resulted in a generalization of an operational definition to become:

(1) What do you want to accomplish?

(2) By what method will you accomplish it?

(3) How will you know when you have accomplished it?

Whatever you may be doing, until you can answer all three of these questions, you do not have an operational definition but merely a basis for an argument. Shewhart understood this, and in his work he used the concept of an operational definition in the development of the "operation of statistical control." In fact, a process behavior chart, what Shewhart called a control chart, is an operational definition, as Shewhart explains in the opening paragraph of *Statistical Method*:

> "Corresponding to these three steps there are three senses in which statistical control [i.e. process behavior charts] may play an important part in attaining uniformity in the quality of manufactured product: (a) as a concept of a statistical state constituting a limit to which one may hope to go in improving the uniformity of quality; (b) as an operation or technique of attaining uniformity; and (c) as a judgment."

On page 25 of *Economic Control of Quality of Manufactured Product* Shewhart had already written that "This state of control appears to be, in general, a kind of limit to which we may expect to go economically in finding and removing causes of variability without changing a major portion of the…process." Thus, we see that Shewhart was focused on how to operate a process economically with maximum uniformity. To paraphrase this, we could say that the process behavior chart is an operational definition of how to get the most out of any process.

With regard to the first question we might respond that we would like to operate our process up to its full potential. But what is its full potential? The three-sigma limits of a process behavior chart characterize the potential of your process. They define what a predictable process will do, and they approximate what an unpredictable process can be made to do. These limits approximate the *ideal* of what your process can achieve when it is operated with maximum consistency.

But what methodology will allow us to operate at full potential? The running record on the process behavior chart displays the actual process performance. By highlighting those exceptional values where the process performance is inconsistent with the process potential the process behavior chart gives you points to investigate. As you take advantage of these opportunities you can move your process closer to its full potential. Thus the process behavior chart provides you with a *procedure* that you can use to improve your process.

And how will you know when you are operating at full potential? The combination of both the process potential and the process performance on a single chart allows you to make a *judgment* about how close to the ideal your process is being operated.

Hence, the process behavior chart is an operational definition of how to get the most out of your process. It approximates the ideal, provides a method of attaining that ideal, and gives a way to judge how close you have come to that ideal. It is instructive to compare this complete package with other approaches to improvement.

What It Takes for Improvement

For any given product characteristic or process outcome, there will be dozens, if not hundreds, of cause-and-effect relationships which affect that one characteristic. Any attempt at improvement will require that we address this list of cause-and-effect relationships. Fortunately, in order to produce a consistent product stream, we do not usually need to control all of these cause-and-effect relationships. This is because each of the causes will not all have the same effect upon the product characteristic. Some causes will result in large amounts of variation in the product characteristic, while other causes will result in small amounts of variation in the product characteristic. Consequently, our typical model for systems of cause-and-effect relationships is the Pareto principle.

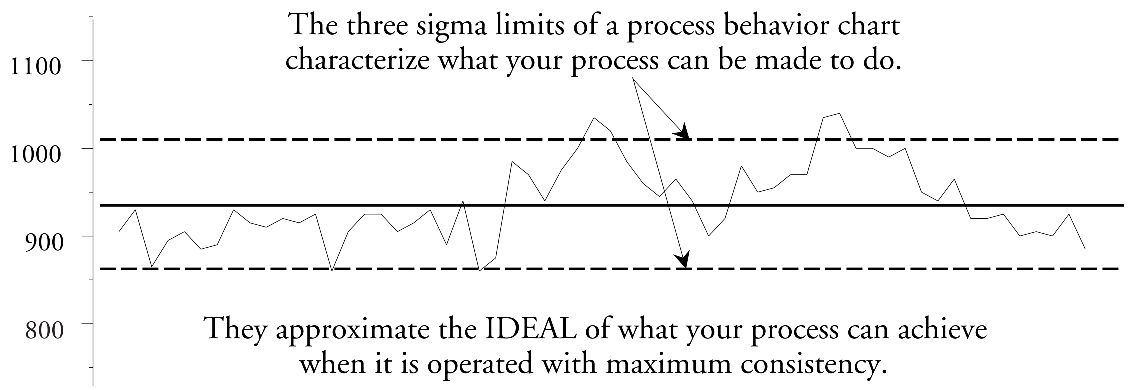

Figure 1: What do you want to accomplish?

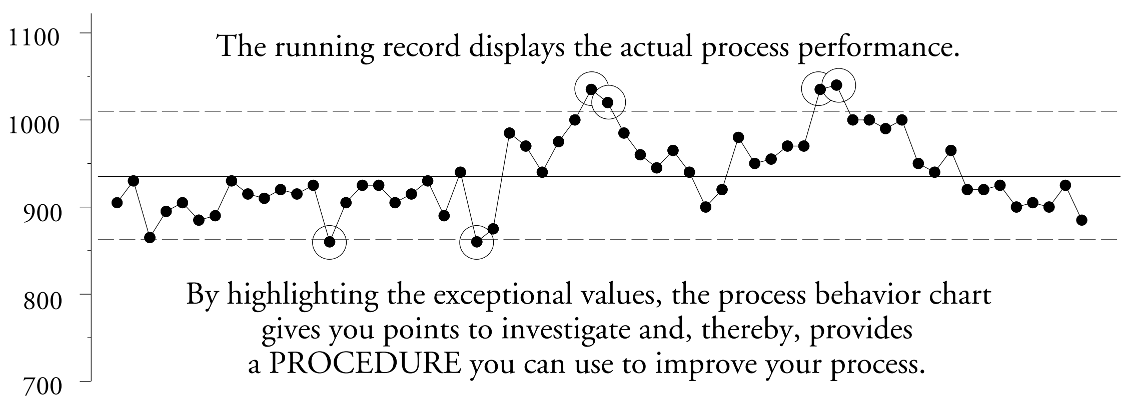

Figure 2: By what method?

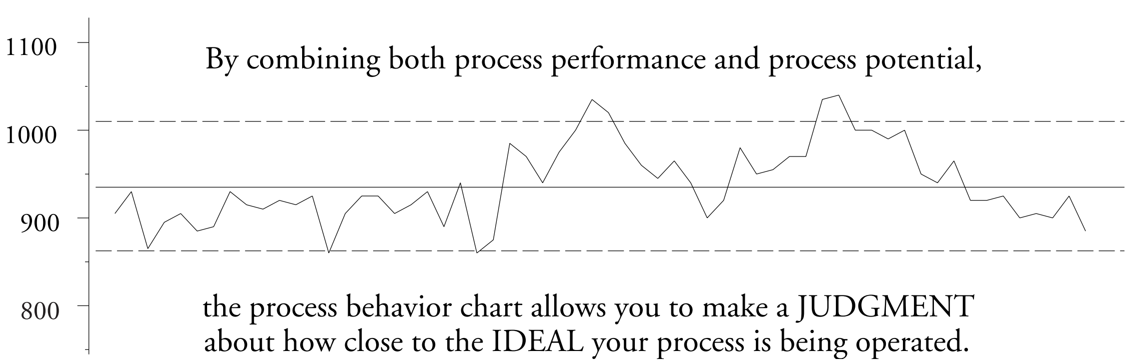

Figure 3: How will you know?

This tendency of Cause Systems to satisfy the Pareto principle does simplify the complex problems of production. In order to make a product we will need to select and control only those factors having a dominant effect upon a given characteristic. We can then ignore the remaining factors. While these remaining lesser causes will create some small amount of product variation, it will be negligible if we correctly identify the dominant causes. In Figure 4, the first four factors account for 78% of the total impact of all 21 factors.

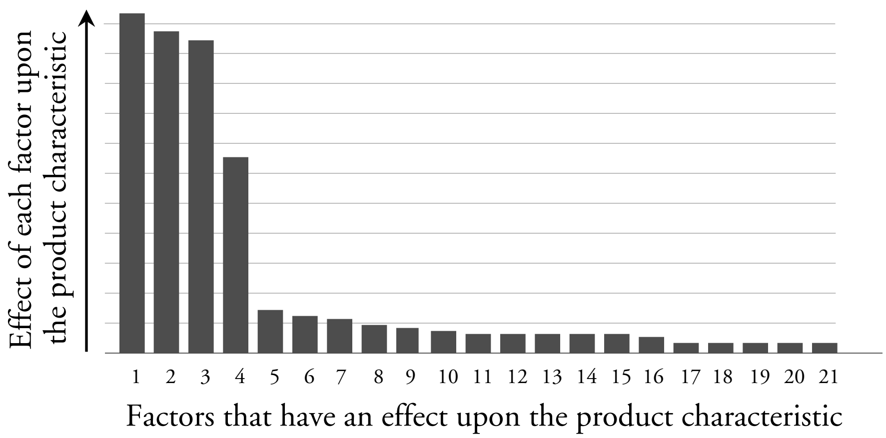

Figure 4: An Underlying Cause System

However, once you have taken care of the dominant causes, you will quickly reach a point of diminishing returns. While the remaining causes will have lesser effects, the effort to nullify these lesser causes will usually be on a par with the effort to nullify the dominant causes. Thus, at some point, it is no longer worth the effort to continue to counteract the lesser causes. In Figure 4, while the first four factors account for 78% of the variation, Factors 5, 6, 7, & 8 account for only 9% of the variation. Because of these diminishing returns we need to have some way to properly separate the dominant causes from the lesser causes. Controlling the dominant causes will have a high payback, while efforts to control the lesser causes will have a low payback.

So how do we identify these high-payback causes? In an experimental approach it is usually done by a combination of guesswork, experience, and research. The problem is to identify those cause-and-effect relationships that we need to control in practice, so we begin with a list of the known and expected causes. Since this list will usually include more causes than we can possibly investigate, we will trim this list by *ranking* these causes according to what we *think* their impact will be and *discarding* those causes thought to have the smaller effects. Then we carry out experiments with the reduced list in order to (a) identify those causes with the dominant effects and (b) determine which levels of these dominant causes will yield a product with the desired characteristic.

The end result of this process is a set of *control factors*—those causes which we think have the greatest impact upon our product characteristic. In production we will hold the levels of these control factors constant. At the same time, we will, of necessity, ignore the remaining cause-and-effect relationships—after all, we did not find them to have dominant effects, so they are relegated to the group of lesser causes.

Figure 5 shows a list of 21 factors that were thought to have an impact upon Product Characteristic No. 2. These 21 factors had been arranged in what was thought to be the order of descending impact.

Since 21 factors were too many for R&D to consider, they decided to evaluate the top ten factors.

While all 10 factors had an effect upon this characteristic, their effects were not all the same size. Factors 5, 1, and 7 were found to have the dominant effects. Thus, Engineering told Manufacturing that they would need to carefully control the levels of Factors 5, 1, and 7 in production. Except possibly for Factor 4, all other factors were thought to have a minimal impact upon this characteristic and, therefore, could be ignored in production.

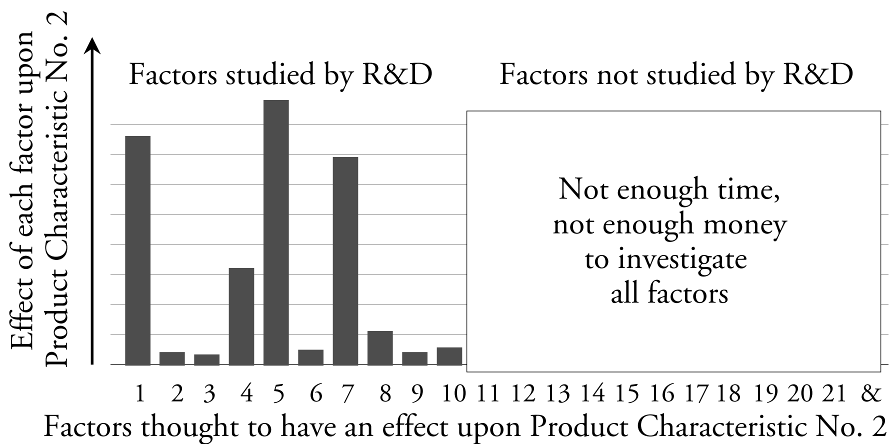

Figure 5: Experiments Always Exclude Some Factors

The production process was then set up using Factors 5, 1, and 7 as control factors for Product Characteristic No. 2. At the start of production they immediately had problems with too much variation in Product Characteristic No. 2. Because this resulted in a high scrap rate they decided to add Factor 4 to the set of control factors. It didn't help. As they fell further and further behind the production schedule, and as the mountain of scrap increased, they began to talk about the "skill" that it took to make this product. Words like "art" and "magic" were used. Inspection and rework facilities were expanded, and soon the production department had settled down to what Deming called the "Western approach to production: burn the toast and scrape it."

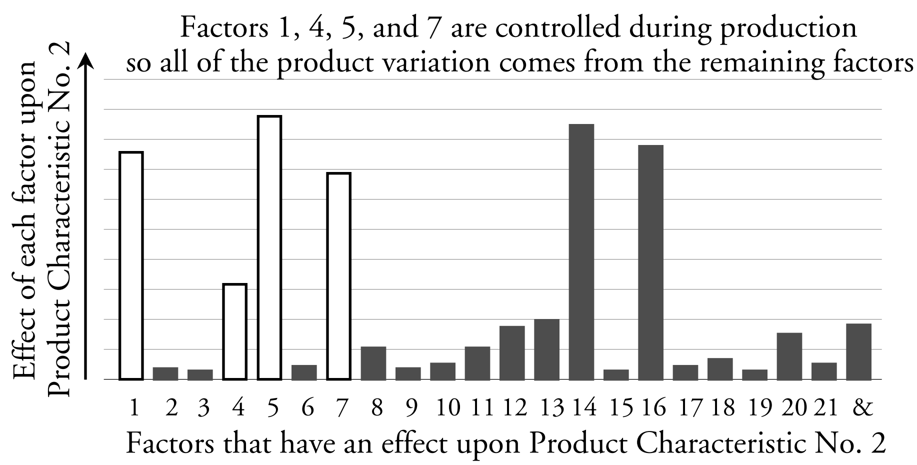

Figure 6: Why can't we make good stuff?

The most common reason for this scenario is seen in Figure 6. While some of the dominant factors were properly identified, others were missed. While the manufacturer was unaware of the impact of Factors 14 and 16, the process continued to be under their influence. Since Factors 14 and 16 had not been studied, and were thought to be part of the lesser causes, the manufacturer was not exerting any control over the levels of these factors. Yet, in the course of events, when the levels of either one of these factors changed it would cause a corresponding change in the product characteristic. While the manufacturer remained unaware of Factors 14 and 16, he suffered the consequences of their effects.

Here, the problem does not come from an inability to control Factors 5, 1, and 7. It is instead due to the fact that Factors 14 and 16 are not part of the set of control factors. Shewhart called these dominant but uncontrolled factors Assignable Causes. Deming called them Special Causes. The lesser causes of Figure 7 were called Chance Causes by Shewhart and Common Causes by Deming.

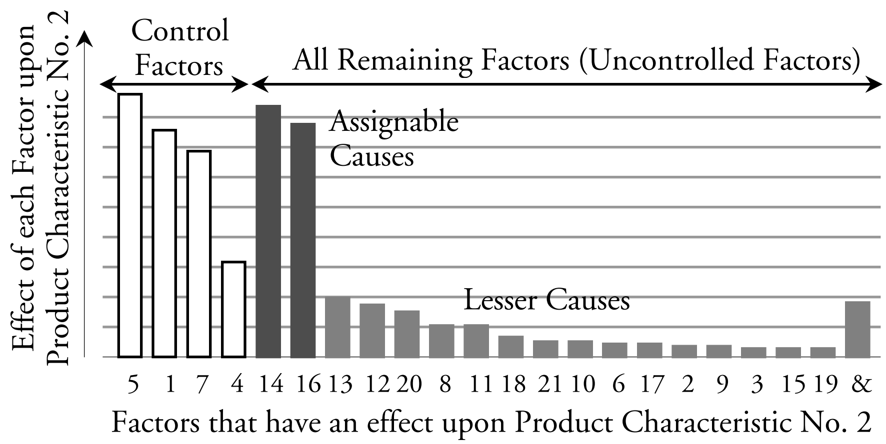

Figure 7: Three Categories of Factors

Thus we have Control Factors, Assignable Causes, and Lesser or Common Causes. Since the levels of the Control Factors are fixed, they contribute little or no variation to the product characteristic. When we place a factor in the control group, we essentially remove it as a source of process variation. *Therefore, effort spent trying to fine tune the Control Factors will be of marginal benefit as long as there are Assignable Causes present.* It does not matter what levels of Factors 5, 1, and 7 we choose as long as we are doing nothing about Factors 14 and 16! ***You cannot optimize any system when some of the dominant cause-and-effect relationships remain unidentified.*** So while experimental studies can be used to identify dominant cause-and-effect relationships, their effectiveness depends on the choices illustrated in Figure 5. For this reason, any approach to improvement that is based *solely* upon experimental studies is inherently flawed and incomplete.

Since the group of Assignable Causes will contain *all* of those dominant causes that are not in the set of Control Factors, this group will be the source of *most* of the unexplained variation in the product. In Figure 7, Factors 14 and 16 account for almost 60 percent of the remaining variation. Assignable Causes are those factors that give managers gray hair and ulcers. This group is the major contributor to excess costs of production, low quality, scrap, and rework. Therefore, effort spent in identifying Assignable Causes and making them part of the group of Control Factors will generally have a very high payback.

Finally the group of Common Causes will be the source of the run-of-the-mill, routine variation that is always part of the background of all production processes. Effort spent trying to control the Common Causes will, at best, yield small returns, and will usually be effort wasted.

It is important to note that whenever we identify a dominant cause-and-effect relationship, and then make that cause a member of the set of Control Factors, we will remove a source of variation from the process. This means that the dominant causes discovered in the R&D phase will result in reduced product variation. However, a partial understanding of which factors are dominant will only result in a partial reduction of variation. And every R&D effort is limited by the ability of the researchers to identify the key factors *in advance*. Whenever we decide to conduct an experiment there will be some factors that we choose to study, and there will be other factors that we choose to leave out of the study. These excluded factors may be held constant, or randomized, or ignored, but since they are not studied their impact remains unknown. Whenever a dominant cause ends up being held constant, or randomized, or ignored, your experiment can only give you a limited and partial understanding of your process.

However, in spite of your *understanding* of which factors are dominant, your *process* will always be subject to the effects of *all* of the Uncontrolled Factors. When the set of Uncontrolled Factors contains Assignable Causes your process will suffer the consequences. Therefore, since Experimental Studies will always be limited in scope, we will also need the ability to conduct Observational Studies.

Observational and Experimental Studies

Observational studies are studies where the data are obtained as a by-product of some ongoing operation. These data may be deliberately and intentionally collected, but they are still a by-product of some process while that process is being operated in an ordinary manner. In other words, in an observational study, the data track the process.

On the other hand, experimental data are collected under special conditions where those different conditions are created for the express purpose of obtaining the data. Experimental studies will always result in a fixed amount of data, collected under *different* conditions, while observational studies will result in ongoing streams of production data, usually collected while the known inputs (the control factors) are held constant.

Because experimental data are collected under special conditions, they tend to be more expensive than observational data. Moreover, because of the way they are obtained, we expect experimental data to represent the differences between the special conditions. Thus, when we analyze experimental data we are looking for differences that we have paid good money to create and that we believe are contained within the data. Moreover, the fact that we will have to conduct more experiments if we can not find the expected differences will tend to make us choose a less conservative, and more exploratory, analysis for our experimental data.

Table 1: Observational Studies vs. Experimental Studies

Observational Studies	Experimental Studies
Additional Data Available	Fixed Amount of Data
One Condition Present	Two or More Conditions Present
Should Be No Signals	Should Be Some Signals
Sequential Analysis Procedure	All Data Analyzed at One Time
Conservative Analysis Used	Traditional or Exploratory Analysis Used

When analyzing data from an observational study, the fact that the data were supposedly collected under one condition will mean that we do not expect to find any differences within the data. Furthermore, since any differences that do occur will often indicate unplanned changes in the process, we will want to be sure about any differences we find. Since additional data will usually be available, we can afford to play a waiting game with observational studies. The limits on a process behavior chart provide a conservative analysis for each new point added to the chart. Therefore, when a point goes outside these three-sigma limits we will have strong evidence that the process has changed before we take action.

As seen in Figure 8 there are several different analysis techniques that can be used with experimental studies, while some techniques may be used with either type of study.

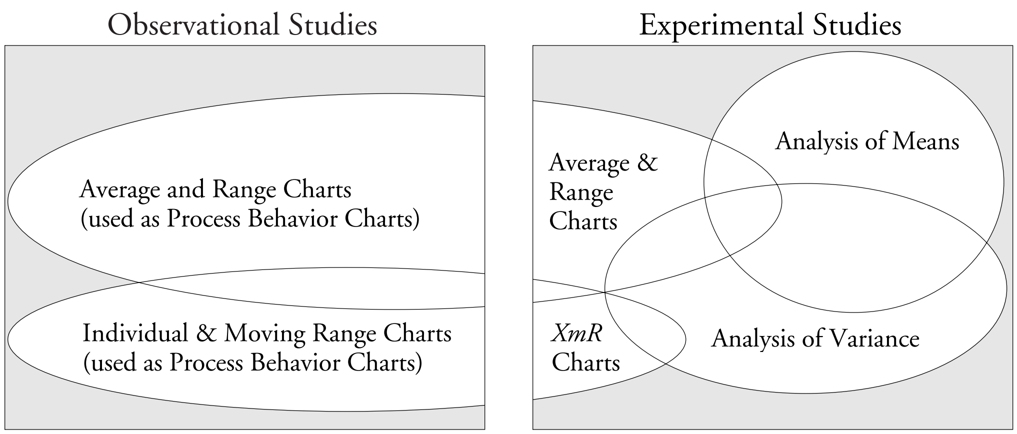

Figure 8: Analysis Techniques for Observational and Experimental Studies

The Individual Value and Moving Range Chart (*XmR* Chart) and the Average and Range Chart were created for use with observational studies. When they are used in this way they may be said to be "Process Behavior Charts." They allow us to identify Assignable Causes of exceptional variation such as Factors 14 and 16, so that we can move them from the set of Uncontrolled Factors to the set of Control Factors, and thereby reduce the variation in the product stream. Thus, in order to be effective, process improvement efforts will require the use of observational studies.

Where Does Six Sigma Fit In?

There are many different programs being taught and used today under the general heading of "Six-Sigma." While these programs vary widely in their content there are some common elements that are found in most. First among these is the idea of using improvement projects to achieve breakthroughs to new levels of quality and savings. Implicit in this approach is the idea that every process needs to be reengineered. To this end experiments are conducted to determine what new materials, new procedures, or new technologies to use. Thus, the six sigma approach is primarily based upon experimental studies.

When adding new elements to a process as part of a process upgrade, or when designing a new process, experimental studies allow you to examine the impact of selected factors upon a product characteristic. However, our model for the nature of data tells us that trusting in the results of a series of experimental studies can never be completely satisfactory. Experiments are very successful at answering specific questions and confirming suspected relationships. They are of limited utility when we do not know what questions to ask.

The very nature of an experiment demands that while we study some relationships we will also have to ignore other relationships. For this reason, what we learn from experiments is likely to be true, but it is also likely to be incomplete. Thus, our model for the nature of data tells us that we need to complement and complete the results of any experimental study by using an observational study. For it is only with an observational study that we learn what our process does while *all* of the cause-and-effect relationships are present. This is the only reliable way to determine if our process contains any Assignable Causes like Factors 14 and 16. Observational studies are a necessary step in learning from the process data and in developing a complete understanding of our processes.

The third, and perhaps the most important, implication of our model for the nature of data is that we should start any investigation of an existing process with an observational study. If we try to adjust, improve, or optimize an existing process without first checking for the presence of Assignable Causes we may end up playing with less than a full deck. As a result we may find ourselves erroneously experimenting with the levels of Factor 4 rather than working to remove the effects of Factors 14 and 16.

So while the DMAIC models of six sigma emphasize the experimental approach to improvement, there is an alternate route that also needs to be part of your improvement strategy. No DMAIC model that this author has seen considers what can be done by operating the current process up to its full potential. In fact, most six sigma models do not make any distinction between processes that are operated predictably and processes that are operated unpredictably. This blind spot can be remedied by using process behavior charts at all phases of the DMAIC model. Operating a process predictably is an achievement. It requires operational discipline. If you are not operating your current process predictably, you are unlikely to operate an upgraded process predictably. And the only way to get the most out of any process is to operate it predictably.

All DMAIC models are expanded versions of the idea of an operational definition, and as we have seen, the process behavior chart is itself an operational definition of how to get the most out of any process. In short, there is nothing wrong with a DMAIC model that cannot be remedied by using process behavior charts at each step in the DMAIC model. Process behavior charts will help you to *define* which processes are not operating up to their full potential and are therefore in need of improvement. Process behavior charts automatically filter out the uncertainty introduced by the *measurement* process. Process behavior charts provide a basis for *analyzing* how a process is operating. Process behavior charts identify opportunities for process *improvements*. And process behavior charts give us a way to continue to operate a process up to its full potential in the future (i.e. to *control* it). Hence we can transform the complexity of the DMAIC models into a very simple and straightforward process of continual improvement by using process behavior charts at each step.

Summary

Given that we now have two different ways to identify the dominant cause-and-effect relationships in our process, should we start with experimental studies or observational studies?

While it is possible to experiment with unpredictable processes, it is complex and difficult to do so. Experimentation is much more effective when all of the dominant factors are included in the study. Therefore, the simplest, easiest approach to process improvement with an existing process is to use an observational study. By taking advantage of the ability of an observational study to identify the unknown but dominant factors that are lurking around most processes you can quickly improve both the process and the process outcomes. Once you have your process operating up to its full potential you may well find that further improvements are not needed. By taking advantage of the opportunities presented by a process behavior chart it is possible to cut the process variation in half, or even more. Shewhart provided examples of this in *Economic Control of Quality*. My own clients have given me examples where the process variation was reduced to one-third, one-fourth, and even one-fifth of what it had been originally. The observational approach has a proven track record.

However, if you should need to further improve a process that you are currently operating at maximum consistency, then experimental studies can be used to determine how to reengineer the process. Likewise, when designing a new process experimental studies are appropriate and necessary. Thus, both observational studies and experimental studies are needed in any effective program of process improvement. Do not try to do the job using only one or the other.