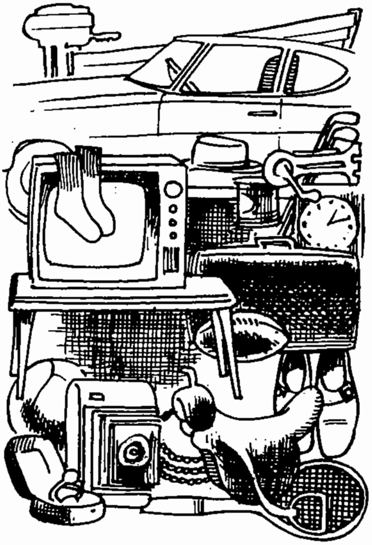

MAKING THINGS RIGHT

W. Edwards Deming *Consultant in Statistical Surveys*

WHAT IS THE STATISTICAL CONTROL OF QUALITY?

The statistical control of quality is the use of statistical methods in all stages of production—in design of product, in tests of product in the laboratory, in tests in service, for specifications and tests of incoming materials and assemblies, and for achieving economies in production, maintenance, and replacement of machinery and equipment, economies in inventory of parts for repairs of machinery, even economies in inventory to meet predicted demand.

Inspection is a very important function in production. The effects of instruments, machines, and human observations jointly create figures that must be transcribed onto forms constructed for the purpose. Faults recorded in inspection may be inherent to the product, or they may be caused by faulty instruments or gauges, or even by poor measuring practice.

We must be content in this article to limit ourselves to a few simple examples of statistical control of quality drawn from the production line. In the first two examples the aim will be to detect the existence of special causes of trouble, for the operator to correct. In the third example the aim will be to measure the effects of common (environmental) causes of trouble, for management to correct. In the real world, we are always working on both kinds of causes. We hope the reader will see in the examples the distinction between special causes and common causes and how they affect the variability of the process or lead to other kinds of trouble.

EXAMPLE 1: FUDGING THE DATA

Figure 1 shows the distribution of diameters in centimeters, these being the results of the inspection of 500 steel rods. Such a graphic representation of a distribution is called a histogram. The lower specification limit (abbreviated LSL) of the diameter of these rods was 1 centimeter. Rods smaller than 1 cm. would be too loose in their bearings, and such rods would be thrown out (rejected) in a later operation, when they must be fitted to a hole. Rejection means loss of all the labor that was expended on the rod up to this point, as well as loss of material and of overhead expense.

The horizontal axis in Figure 1 shows the centers of intervals of measurements; for example, 0.998 stands for rods that measured between 0.9975 and 0.9985 cm. The vertical axis is labeled to show the number of rods that fell into an interval of 0.001 cm. on the horizontal axis. For example, about 30 rods were in the interval centered at 0.998. It appears from the distribution that $10 + 30 + 0 = 40$ rods failed because they were too small.

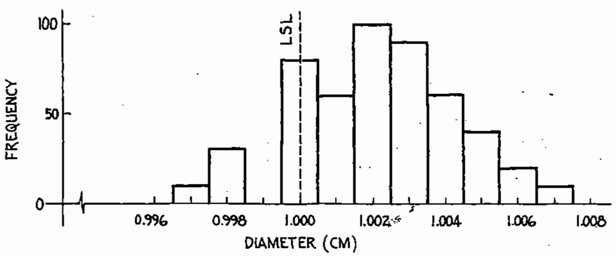

FIGURE 1

Distribution of measurements on the inside diameters of 500 steel rods. The chart detected the existence of a special cause of variation, a fault in recording results of inspection

A distribution is one of the most important statistical tools, when used with skill, yet it is extremely simple to construct and to understand.

Figure 1 is trying to tell us something. The peak at just 1 cm. with a gap at 0.999 seems strange. It looks as if the inspectors were passing parts that were barely below the lower specification, recording them in the interval centered at 1.000. When the inspectors were asked about this possibility, they readily admitted that they were passing parts that were barely defective. They were unaware of the importance of their job, and unaware of the trouble that an undersized diameter would cause later on.

This simple chart thus detected a special cause of trouble. The inspectors themselves could correct the fault. When the inspectors in the future recorded their results more faithfully, the gap at 0.999 filled up. The number of defective rods turned out to be much bigger, 105 in the next 500, instead of the false figure of $10 + 30 + 0 = 40$ in Figure 1.

The results of inspection, when corrected, led to recognition of a fundamental fault in production; the setting of the machine was wrong. It was producing an inordinate number of rods of diameter below the lower specification limit. When the setting was corrected and the inspection carried out properly, most of the trouble disappeared.

The upper specification limit had its problems also, but they were not so serious. A rod that is too large in diameter can be tooled off to fit. This is not the economic way to achieve the right dimension, but it is cheaper than to lose all the labor expended up to that time on the rod. The next problem was accordingly to increase uniformity and work on the correct centering of the average diameter, to reduce the number of defectives with wrong diameters.

EXAMPLE 2: DETECTING A TREND

The second example deals with a test of coil springs one after another as they come off the production line. These springs are used in cameras of a certain type. According to the specifications, the spring should lengthen by 0.001 cm. for each gram of pull. These springs are relatively expensive, and are supposedly made to exacting requirements. The length of any horizontal bar in the histogram at the right in Figure 2 shows how many springs the inspectors recorded with the elongation shown. We have turned it sidewise for convenience. This histogram represents measurements on 50 springs manufactured in succession. It will be noted that the distribution is symmetrical and is centered close to the specification; furthermore all 50 springs were within the upper and lower specification limits. One might be tempted to conclude from this histogram alone that the production of this spring presents no problems. However, another simple but powerful statistical tool, called a *run chart,* indicates trouble, as we now explain.

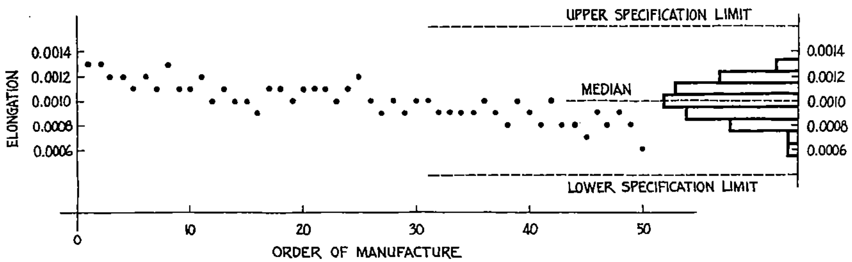

FIGURE 2

Run chart for 50 springs tested in order of manufacture. The chart shows a definite trend downward and thus reveals the existence of a special cause of variation, which it is important to correct. The frequency distribution alone could not detect this trouble

A run chart is merely a running record of the results of inspection. The horizontal scale shows the order of the item as produced, and the vertical axis shows the measurement for that item. In Figure 2, the elongations of the 50 successive springs are plotted on the vertical scale. A run chart has several simple uses. For example,

(1) A run of six or seven consecutive points lying all above or all below the median—the middle point in height—signifies with near certainty the existence of a special cause of variation, usually a trend.
(2) A run of six or seven points successively progressing upward or successively progressing downward has the same significance.

In no instance in Figure 2 is there a run up or a run down of length 6. It so happens that the median of the 50 points falls midway between the upper and lower specification limits. This would be good, but we note that the opening burst of points at the left of the figure has 10 points in succession above the median. Fifteen out of 18 points after point 29 fall below the median. These observations give a statistical foundation for the conclusion that, although the points vary up and down, there is a general drift downward. You may feel that your eye was good enough to detect this trend without knowing from theory that a run must have six or seven points above or below the median to detect with near certainty the existence of trouble, and in this example, you would be correct, but in more complicated examples such trends are often not detectable by eye.

Knowledge of what lengths of runs are required to indicate trouble is also valuable but secondary in problems of production. Indeed, it is an important statistical point that some of the most powerful statistical techniques are simple, as in our examples here. It was their widespread use, which began about 1942, that laid the foundation for the statistical control of quality, which of course has since grown into all phases of management. This movement led to the organization years ago of the American Society for Quality Control, over 23,000 strong in 1970.

In our camera-spring example, either the production process is in trouble or the apparatus used for testing is giving false readings. Correction is vital, whatever be the source of the trend. If it is the tension of the spring that is drifting downward (and not the testing apparatus), defective springs will be produced in the immediate future. If the source of the trend is faulty testing, then the tests are misleading, and may have been giving faulty reports on all the springs produced recently.

In this particular case, the trouble lay in a thermocouple that permitted the temperature to drift during the annealing of the springs. The process was headed for trouble. The simple run chart detected the trend before trouble occurred. The operator himself, seeing the trend, was able to head off trouble.

The reader may note that the histogram and the run chart in Figure 2 were plotted from the same data, yet they tell different stories. The histogram by itself gives no indication of anything wrong; it could have indicated unsatisfactory positioning. The run chart, however, leads us to suspect the existence of something wrong, a trend that, unless corrected, would soon lead to the production of defective springs.

It is interesting to note that if the points in Figure 2 had been plotted in random order instead of one after another in the order of production (1, 2, 3, and onward to 50), the run chart would have lost its power to detect a trend. Statisticians are thus not only concerned with figures, but with the relevant figures. In this instance, the order of production was relevant—very relevant—and was used to make the run chart. The histograms in Figures 1 and 2, on the other hand, do not make use of the order of production. They would remain unchanged, regardless of order: they depend only on the numbers recorded as results of inspection. The histogram in Figure 1 nevertheless did its work; it told us that something was wrong (namely, in the inspection itself). A run chart in connection with Figure 1 would not have added any relevant information. The histogram in Figure 2, however, was helpless to detect the existence of anything wrong. Judging by it alone, without the run chart, we could not have detected impending trouble.

EXAMPLE 3: MEASUREMENT OF COMMON (ENVIRONMENTAL) CAUSES

The first two examples dealt with special causes, specific to a designated worker or to a machine or to a specific group of workers. Statistical techniques point to specific sources of trouble when the process is nonrandom. The same statistical methods also tell the worker to leave things alone, to avoid overadjusting when attempts at adjustments would be ineffective or cause even greater variation than now exists.

There is another kind of problem that faces the management of any concern. No matter how skilled the workers, and no matter how conscientious, there will be at least a bedrock minimum amount of trouble in production owing to common or environmental causes. All the workers in a section work under certain conditions fixed by the management, or one might say, by the environment, which only the management can alter. For example, all the workers use the same type of machine or instrument. They are all doing about the same thing, and are using the same raw materials (which might be semifinished assemblies). They must put up with the same amount of noise and smoke.

It used to be supposed by management that all troubles came from the workers: that if the workers would only carry out with care the prescribed motions (soldering a joint, placing a part, turning a screw), then the product would be right with no defectives. This kind of reasoning does not solve the problem. Alert management can look into the problem with "infrared vision," supplied by statistical techniques.

An example was a small factory that made men's shoes. The machinery that sews soles is expensive. Time that an operator spends rethreading the machine and adjusting the tension after a break in the thread is time lost. Minutes lost may add up to hours and even days in the course of a month. There is not only the loss of rent paid for the machine and wages of the operator, but loss of labor and materials, nonproductivity of floor space, light, and increases in general overhead expenses. In this factory, about 10% of the working time was being spent rethreading the machines and adjusting the tension. Management was rightly worried. The trouble became obvious with a bit of statistical thinking. Observations on the proportion of time lost by the individual workers provided data for a chart similar to Figure 3. This figure showed that all the operators were losing about the same amount of time rethreading their machines. In fact, the time lost per day per man showed a pattern of randomness. This uniformity across operators could only point to the environment. What was the trouble?

The trouble turned out to be the thread. The management of the company was trying to save money by buying second-grade thread that cost 10¢ per spool less than first-grade thread. Penny wise and pound foolish! The

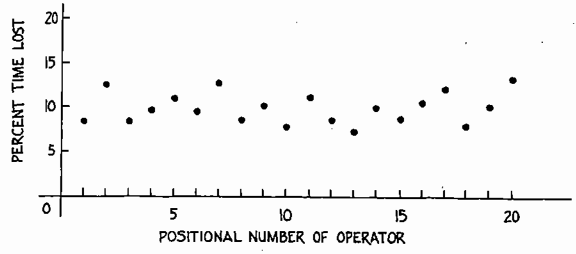

FIGURE 3

Time lost by each of 20 operators

savings on thread were being wiped out and overwhelmed many times over by troubles caused by poor thread.

A change to thread of first grade eliminated 90% of the time lost in rethreading the machines, with savings many times the added cost of better thread.

What is the distinction between this example and examples 1 and 2? In examples 1 and 2, the workers themselves could make the necessary changes, and they did. In this example, the operators were helpless. They could not put in an order for thread of first grade and scrap the bad thread. Their jobs were rigid—work with the materials and machines supplied by the management. They all worked with the same bad thread; that is, they all worked under a common cause of trouble. Management is responsible for common (environmental) causes; therefore only management could change the thread.

But how is management to know that there are common causes of trouble? The answer is simple: common causes are always present. Management needs a better answer, however; management needs a graphical or numerical measure of the magnitude of the trouble wrought by common causes. Without statistical techniques, management can have no accurate idea about the magnitude of the trouble being caused by conditions that only management can change.

Charts such as Figure 3 tell the management that there is a problem, that the time lost on rethreading will not go below 10% until management makes some fundamental change. The change in thread in example 3 was a fundamental change. What to change is not always as easy to perceive as it was in this example, however. Sometimes a series of experiments is required to discover the main causes of trouble. Statistically designed experiments have led to the identification of common causes such as raw materials not suited to the requirements, poor instruction and poor supervision (almost synonymous with unfortunate working relationships between foremen and production workers), and vibration.

Shift of management's emphasis from quantity to quality is one common environmental cause of trouble. The production workers continue to work with emphasis on quantity, not quality. Discussion of methods by which management may direct the shift from quantity to quality, however important, is beyond the scope of this essay.