Center for Quality and Productivity Improvement
University of Wisconsin
610 Walnut Street
Madison, Wisconsin 53705

(608) 263-2520
(608) 263-1425 FAX
quality@engr.wisc.edu

Report No. 158

The Role of Scientific Problem Solving and Statistics in Quality Improvement: Some Perspectives

Søren Bisgaard, Ph.D.

March 1997

The Center for Quality and Productivity Improvement cares about your reactions to our reports. Please direct comments (general or specific) to: Reports Editor, Center for Quality and Productivity Improvement, 610 Walnut Street, Madison, WI 53705; (608) 263-2520. All comments will be forwarded to the author(s).

CQPI
University of Wisconsin

Report No. 158

Report Series
in Quality and Productivity

The Role of Scientific Problem Solving and Statistics in Quality Improvement: Some Perspectives*

Søren Bisgaard

Director, Center for Quality and Productivity Improvement and
Associate Professor, Department of Industrial Engineering

Univeristy of Wisconsin–Madison, USA

ABSTRACT

Scientific method of observation and experimentation play a key role in quality improvement. In this article, I provide numerous examples of the use of scientific method and argue that this is a vital catalyst for Total Quality Management. Perspectives for the future are also provided.

---

*Invited plenary talk, presented at the European Quality Congress, Trondheim, Norway, June 1997.

CQPI Report No. 158, March 1997

The Role of Scientific Problem Solving and Statistics in Quality Improvement: Some Perspectives

Søren Bisgaard, Ph.D.

Scientific method of observation and experimentation play a key role in quality improvement. In this article, I provide numerous examples of the use of scientific method and argue that this is a vital catalyst for Total Quality Management. Perspectives for the future are also provided.

INTRODUCTION

Quality control and quality improvement are historically by most considered relatively new concepts. For its conception we may refer to Dr. Walter Shewhart's famous control chart memorandum from 1924, or his path breaking book, "The Economic Control of Quality of Manufactured Products", from 1931. In this book, and many of his later publications, most notably the small 1939 book "Statistical Method from the Viewpoint of Quality Control", Shewhart made it clear that to him the guiding principle behind quality control and improvement was the use of scientific method of observation, experimentation and, in particular, statistics.

The idea of using scientific approaches to create and improve products and services that satisfy customers is, however, an idea that easily can be traced back to the beginning of the Industrial Revolution, if not much earlier. It is a universal idea, and as I see it, the modern quality movement is a natural outgrowth and extension of the 18th century Industrial Revolution, itself spurred on by the Enlightenment and the earlier Scientific Revolution, see Bronowski and Mazlish (1960). For example, when Josiah Wedgwood and his friends in the Lunar Society of Birmingham, an informal group of industrialists and scientists, meet once a month at full moon, hence the name, the purpose was to discuss the latest scientific ideas. Of particular interest appear to have been how scientific method itself could be utilized to solve the many practical problems involved in developing new, better and cheaper products and reduce waste. In particular, Josiah Wedgwood, a potter by trade, developed a high temperature pyrometer so that he, in his extensive experimental effort to develop high quality ceramic, could better measure and control the effect of temperature. For this achievement he was elected Fellow of the Royal Society of London, the oldest most prestigious scientific society in the World. Wedgwood's meticulous and systematic experiments, some of which are preserved to this date, will impress any modern industrialist. Moreover, his organization of factory work at his Etruria Works is an early example of the use of quality management principles, see Tames (1979).

What these early industrialists have in common with 'today's quality professionals is the idea of using scientific method to solve problems and develop new products and services that satisfy their customers. Many things have clearly changed since the 18th century. Even the 1930s, when Shewhart and other quality pioneers laid the foundation for our field, may seem a long time ago. Many things will continue to change in the future. However, I am convinced that one powerful idea will prevail and that is the use of scientific method to everything we do. Scientific method is not all of Total Quality Management (TQM), but is clearly a catalyst without which Total Quality Management will not succeed. Moreover, the incorporation of scientific method as an integral part of its theoretical foundation is the one distinct concept that sets TQM apart from many popular management fads.

Scientific method means different things to different people and is clearly a large conglomeration of concepts. To some it might mean the use of sophisticated equipment in fancy laboratories applied to develop futuristic concepts, or to the advancement of fundamental theories of nuclear physics, biology or chemistry. What I have in mind and want to discuss in this article is, however, not related to any particular equipment or application area, and not confined to fundamental research. I want to focus on the idea itself and the process of a scientific investigation — the efficient, effective and systematic process of learning and discovering. This is whatever process a scientist, or for that matter any intelligent human being, uses when trying to find out something, or anything, that we don't know. It is the process of systematic discovery.

Even this more philosophical concept is not without some controversy as to what it is, and obviously will depend to some degree on the particular application. However, to most people scientific method involves observation and experimentation as key concepts. The Scientific Revolution gained momentum in the 16th century when the idea began to germinate that physical phenomena could be explained by cause-and-effect relations as opposed to being caused by divine intervention. It is difficult to appreciate today what a profound thought revolution that must have been. The triumph of 17th century astronomy and physics was the possibility of explaining the movements of the planets by simple mechanical cause-and-effect relationships similar to those of a clockwork. This had profound implications for the development of our modern World view. It facilitated the widening belief in cause-and-effect explanations as opposed to dogma. For example, a brewer's batch of bad beer was caused by something physical rather than metaphysical. If a cause-and-effect relationship could be established, the process could be controlled and improved. With that idea scientists and industrialists began looking for causes where nobody had looked before. Moreover, it was found, by Francis Bacon in particular, that these cause-and-effect relations could be discovered inductively by careful observation and experimentation rather than through the reading of ancient manuscripts. Astronomy and physics lead the way and other areas such as medicine, chemistry and biology followed with great benefit to mankind.

Soon after the early days of the Scientific Revolution, as already indicated, pragmatic scientists and industrialists got together, primarily in England but also elsewhere, and by mutual inspiration caught on to the idea that scientific method of observation and experimentation could also be used to discover the cause-and-effect relations governing the manufacture of quality consumer products. The manufacture of just about anything involved then, as it does now, much time consuming and complex problem solving and trouble shooting. For example, non-uniformity of clay and lack of control of kiln temperature, cause ceramics to break or look unsatisfactory resulting in scrap and waste, delivery delays, unhappy customers and ultimately business loss. The problem of manufacturing high quality ceramics on a large scale was what Wedgwood elegantly attacked more than 200 years ago with scientific method. What many do not appreciate is that his problems were conceptually not any different from those facing present day high tech manufactures of micro chips for computers. They also struggle to increase yield. They are often just as much in the dark as to what cause yield problems as were Wedgwood and his contemporary potters. The same approach that Wedgwood applied, the use of observational studies and systematic experimentation, applies today to troubleshooting the most sophisticated high-tech industrial quality problems. However, we have a slight edge over our pioneering colleagues from the early days of the Industrial Revolution. We have available a broad range of statistical concepts and tools that can be of help making the process of scientific investigation more efficient and increase the chance of discovering new ways to increase quality and productivity to satisfy the customers.

As indicated above, discovering cause-and-effects relations (assignable causes) using observational studies and experiments is at the heart of the Industrial Revolution, is a key component of the modern quality movement, and will surely be of great help to us in improving the quality of products and services, and for that matter, the quality of any area of human endeavor in the future. To better explain this, I will in the following sections of this article explain and exemplify the role of observational studies, experiments and statistics and reflect on the role of these ideas for quality improvement in the future.

SCIENTIFIC METHOD

It is difficult if not impossible to provide a single formula or definition for what scientific method is. Nevertheless, the term is in common use, and in most cases I think that there is a general consensus that it involves most if not all of the following six elements, not necessarily applied in that order, and clearly used iteratively:

1. Recognition and formulation of the problem.

2. Collection of data from observational studies.

3. Planned experimentation.

4. Generation of a working hypothesis or model by induction from patterns in data, from "in sight", current understanding, hunches, and creative thinking.

5. Making deductions (predictions) from the hypothesis or model.

6. Comparing and testing predictions and deductions from models with data and drawing inferences. If the data confirms the deductions from the hypothesis or model, the hypothesis is more believable and further consequences can be explored. If not, the question of why or what is wrong with the model leads to the generation of new ideas, conjectures and hypotheses.

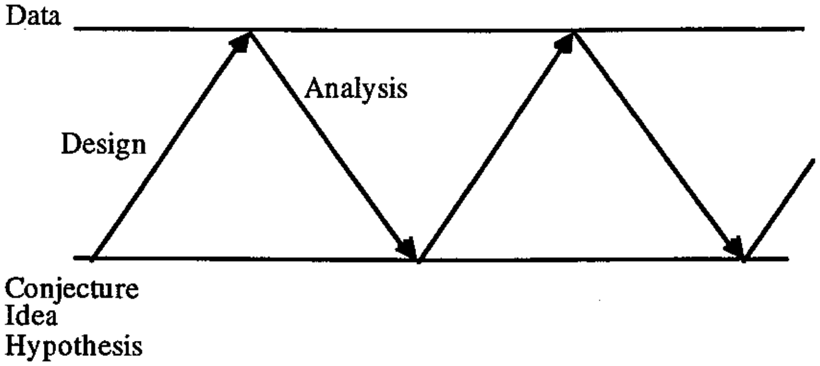

Figure 1. The iterative nature of a scientific investigation as conceptualized by Box (1976).

Shewhart, keen on making the connection between quality improvement and scientific method, used the cycle, now known as the Shewhart cycle or slightly modified as Deming's Plan-Do-Check-Act Cycle, see Deming (1986), and explained that "the three steps [specification, product and control] correspond to the three steps in a dynamic process of acquiring knowledge." Box (1976) used a slightly different graphical depiction of the process of learning and scientific discovery shown in Figure 1. Juran's Quality Trilogy, see Juran (1992), is another conceptualization of the scientific process as it relates to quality control and improvement. Ishikawa, see especially Ishikawa (1985), also shows a strong adherence to the use of scientific method. In fact, I will claim that the unifying idea among all these key contributors to our current conceptualization of TQM is this particular emphasis.

OBSERVATIONAL STUDIES

Philosophers of science, and the philosophy of science literature, appears to me of little help if one is looking for a guide to scientific strategy. Beveridge (1980, pp. 57-58) has said about Karl Popper, considered by many the leading 20th Century scholar on scientific method, that "... [he] does not deal with the origin of hypothesis — surely the very heart of scientific discovery. He says there is no such thing as induction and he dismisses creative thinking as outside his field ... He disowns any obligation to deal with the creative act, with the discovery process itself, yet this is the very thing scientists are striving to do, it is what research is all about!" I tend to agree with Beveridge, and will therefore instead provide examples to illustrate what I mean by scientific discovery and this process's importance for quality improvement in particular.

Example 1. To exemplify the inner working of the process of scientific discovery, let me briefly sketch how scientists were able to develop our present day conception of the Solar System. To start with a definite date, on the 21 of August 1560 a fourteen year old Tycho Brahe experienced a Solar eclipse. (For this example, see Dreyer 1963, Butterfield 1962, and Bronowski and Mazlish, 1960) This event made him devote the rest of his life to astronomy. At the time a conflict had developed about whether Ptolemy's Earth centered or Copernicus' Heliocentric conception of the World was right. The fight was mainly fought over how to interpret old astronomical books handed over from the ancient Greeks. Tycho Brahe's most significant insight was that there was no point in reading old manuscripts, and philosophizing about which system was right, as traditional scholars did. Instead he suggested that the answer could be found by meticulous observation of the movements of the stars and planets. Tycho Brahe then proceeded to do so for more than 30 years. Because of a political dispute he decided late in his life to leave his native Denmark, relocate to Prague, and continue his work there with the help of a young assistant named Kepler. When Tycho Brahe died in 1601, Kepler inherited the-by-then extensive records. Under the growing influence of Copernicus' Heliocentric view, strongly promoted by Galileo, Kepler then empirically showed that the data fitted neatly to a model of planets moving in ellipses around the Sun, and established what we now call Kepler's three laws. These law's discovered by induction from the data, were later communicated to Isaac Newton in England by Edmund Halley. Newton then conjectured the existence of a gravitational force inversely proportional to the square of the distance between heavy bodies. On this hypothesis, using his, for the purpose invented method of calculus, Newton proved mathematically that the planets would follow elliptical patterns. Thus he deduced a model from a simple principle, and the predictions from that model were in agreement with the data collected by Tycho Brahe. This provided in turn further credence to Kepler's model and Newton's idea of universal gravity.

Brief as this summary is, we see here a typical pattern of the process of discovery by careful observational study, induction, empirical model building, physical model building, prediction (deduction) and testing. This is characteristic of the inductive-deductive learning process illustrated in Figure 1. But notice that in this spectacular discovery process, the data came first. This data-based process of discovery of "assignable causes," is at the heart of science as well as scientific quality improvement.

Example 2. Now let us consider an example from quality improvement that illustrates the same inductive-deductive cycle in the discovery of root causes. A company making small outboard motors had for years experienced problems with the die casting of carburetors. The material used was a difficult-to-cast alloy of aluminum resistant to saltwater exposure. Because of the difficult nature of the alloy, the casting engineers had over the years grown accustomed to a defect rate of around 24%. They did not think much about it because the scrap could easily be re-melted and the metal used again. However, it occurred to upper management that a 24% defect rate implied that 24% of their capital equipment and worker's time was tied up in idle production of defects, a considerable loss even if the metal could be recycled. To solve this problem a team of engineers with the help of a consultant started to investigate the problem. At first ordinary control charts based on sampling at fixed intervals were made. They showed an erratic behavior. However, no particular pattern emerged and it was not possible to find any assignable causes. The consultant then suggested a more detailed observational study of the process where, for a short period of time, every part made in the two cavity die casting process was put on the floor in the order produced. By recording the cavity number and the exact production order and whether or not the products were defective or not, cumulative sum plots of the data from each cavity were made. This type of plot, very sensitive to small changes in the level of a process, immediately revealed a sudden increase in the defect rate in the middle of the study. Going back to the casting engineers and the workers on the line with these statistical charts, it was found that the sudden increase in the defect rate coincided with the time when the die casting machine had been re-supplied with about 2 tons of hot liquid aluminum alloy brought in by a large crane. It was then conjectured that the change in hydrostatic pressure caused by refilling the reservoir changed the pressure inside the die casting machine. This change of the casting conditions likely caused a surge in the number of defects, it was surmised. Up to this point the method of filling the reservoir had never been questioned. It was standard operating procedure. This study now revealed that a key assignable cause for the high defect rate might be attributable to how the process was run, and not as assumed because the alloy was difficult to cast. After some more work a different way of supplying raw material was instituted, and the defect rate was dramatically decreased. Obviously, this example does not have as far reaching implications as the story about the Solar System. However, we hope the reader upon reflection will recognize that the process of discovery from patterns in data is strikingly similar.

Example 3. The same manufacture of outboard motors also produced six cylinder motor blocks. Sometimes small inclusions were found in these castings late in the production process after several costly machining operations had been performed. When that happened, it was often possible to fix the defectives by welding on material. This rework process, which had been instituted for years was, as many other processes in the company, subject to routine control chart monitoring. Every month these control charts were collected by the quality control department, and summarized into a report that was forwarded to the office of the Vice President. Nobody appeared to have carefully studied the data, however. Apparently the charts were just glanced at, and as long as no dramatic increase occurred nothing was done and the reports were filed away as the "system" told them to do. On one occasion a chart was given to a statistical consultant who proceeded to take a closer look. Using data from a long period of time, and plotting dots for each defect where they occurred directly on a drawing of the motor block — information that was available from the routine data collection sheets — it was revealed that there was a pattern not previously known to where the defects were located. When the consultant showed this to the casting engineers, they first reacted with astonishment. However, they quickly proceeded to explain that, if that pattern was true, the mold needed modification. The naturally occurring impurities were apparently being trapped in the locations of the mold where the defects occurred. By modifying the risers, little "chimneys" on the mold, it was conjectured that the impurities could be directed out of the mold. After such a modification the defect level was significantly reduced.

The point of this story is that just like one night's observation of the sky will never reveal enough information to understand the intricate system of planets moving in ellipses around the Sun, so will looking at few individual defects not necessarily reveal the complex cause-and-effect relations behind defect formation in manufacturing or service processes. However, by carefully studying patterns of defects over longer periods of time, insight into the underlying mechanism can be gained and permanent fixes suggested. This, in my experience, is not done as often as it ought to be. Mechanical application of statistical quality control techniques is what is most often practiced. This is unfortunately implicitly encouraged by many previous efforts to standardized quality procedures. Moreover, most teaching of statistics emphasizes testing rather than the process of discovery. The pay-back from applying an investigative mind set, looking deeper for patterns and root causes is, however, an immense and untapped resource for continuous improvement.

Example 4. We need not confine scientific quality improvement to manufacturing. It applies as well to service functions, and even in the public sector. Some years ago the City of Madison, Wisconsin had serious problems with its service garage for its police cars, garbage trucks and other vehicles. A Pareto chart of the data from a customer satisfaction survey showed that down time, the time from reporting a defect to the vehicle was fixed and delivered back to the customer, was a major problem. The mechanics proceeded to flowchart the process, and after a two month observational study found that about 80% of the down time was spend waiting — waiting for parts, waiting in the parking lot, etc. — and only about 20% was spend on real work. Further investigations revealed that the inventory for spare parts was a major problem causing delays. In turn this was found to be caused by a City purchasing policy of buying on the lowest bid, causing an extraordinarily large number of different types of vehicles and hence an excessive need for different types of spare parts. It was also found that many of the repairs could be prevented. For example, a truck used for salting roads had not been cleaned after the snow season so it rusted away. This was because the City for many years had tried to "save" money by instituting a "if it ain't broke don't fix it" policy. This and many other findings were carefully documented, many statistical charts were made, and it was all presented to the Mayor. He in turn ordered the inventory system fixed, the policy of buying on lowest bid was revised, and most significantly, a preventive maintenance system was instituted. This eventually led to a significant increase in quality, reduction in down time and improved customer satisfaction, as well as a heighten employee morale and therefore improved productivity at this otherwise disillusioned organization.

As all these examples indicate, observational studies play a key role in scientific investigations as well as in quality improvement. This pivotal role was certainly recognized when Shewhart made the control chart the centerpiece of his original exposition on quality control. Simply put, when causes, unknown factors or lurking variables, are not discovered and their nature not fully understood, they may suddenly change in what may appear to be a random fashion, adversely influence the system and cause catastrophic quality problems. On the other hand, when we understand the cause and effect relationship, what to the uninformed observer may appear random, becomes predictable and we can control, neutralize or remove the influence of these factors. As Francis Bacon wisely said, "knowledge is power."

There are, however, many problems where observational studies are not sufficient or do not provide the information quickly enough. Another potent tool is carefully planned experimentation — the deliberate manipulation of a system to stage the evens rather that waiting for them to happen at their natural pace.

EXPERIMENTS

In an experiment, we deliberately change one or several factors to see what effect these changes might have on a system. Defined this way, experiments must have been used since the beginning of mankind. However, the major breakthrough for the method as a systematic approach to discovering new knowledge, came during the Scientific Revolution in the 16th century. Experiments are in most cases performed iteratively as increased understanding is gained in a continuous learning process. To illustrate this process, let me start with a well-documented set of experiments with profound implications for modern life — the discovery of human flight.

Example 5. The possibility of human flight has fascinated people at least since antiquity. However, the problem was solved by two bicycle mechanics in 1903. As young boys the Wright brothers had, as a gift from their father, got a small toy helicopter that, with the pull of a string, flew a few meters. This fascinated them. However, they also understood the danger of flight experiments. The German flight pioneer Otto Lilienthal had just recently killed himself. Lilienthal had flown many times by jumping off a cliff with wings strapped on to his back, but in his fatal experiment, he lost control and went into a spin. The Wright brothers wanted to learn how to fly but also wanted to survive. From Lilienthal's experiments, especially the fatal one, they knew that the key problem was control. Therefore they started a meticulous series of experiments. In 1900 they built and flew many times a large kite from which they learned a great deal about control. Based on these carefully executed and documented experiments, the two brothers went back home to their bicycle shop and build a glider large enough, they estimated, to lift a person off the ground when held up against a strong head wind. The wingspan of this glider was carefully calculated from an empirical formula for the relationship between wing area and lift, established by Lilienthal. This glider, used in the 1901 experiments, however, was a failure. As it turned out Lilienthal's formula was off by at least one order of magnitude. The Wright brothers then realized that they needed to better understand the fundamental principles of aerodynamics. Thus they went home and built a simple experimental device mounted on the handle bar of a bicycle to investigate for themselves the relationship between wing area and lift. After a few experiments this contraption turned out to be inadequate. They then proceeded to build a wind tunnel and a simple device, basically a balance, for measuring lift placed inside the wind tunnel. With this they executed an impressive series of experiments that eventually led them to establish the correct relationship between wing area and lift. Based on this they built a much larger and improved glider. This was used for experiments in 1902 and turned out to be a huge success. It taught them a great deal about how to control the flight, how to make turns, etc. Based on the 1902 experiments they went on to build the motor propelled flyer that in 1903 won them the honor of being the first human beings to fly an airplane. This sequence of experiments is a tour de force illustration of the process of learning by careful experimentation. These two bicyclemechanics were masters of this process. This is, in my estimation, the single most important reason why they succeeded where so many others, including some highly esteemed scientists, failed.

The Wright brothers, as indicated above, understood extremely well the iterative process of learning from scientific experiments and how to use this process to discover cause and effects relations. On this basis they were able to build a quality flying machine. Since their time the methods of planning experiments have undergone a tremendous change. The general process is essentially the same, but the British statistician Sir R. A. Fisher showed in the 1920's that the method of holding all but one factor fixed, while only varying one, is inefficient. Instead he introduced the method of factorial experiments, randomization, blocking and accompanying statistical methods (e.g. analysis of variance) for the analysis of data from such experiments. Fisher's approach is many times more efficient than traditional one-factor-at-a-time methods, and will allow experimenters to better study complex relations between a large number of factors. This is of particular importance when troubleshooting complex manufacturing processes in an effort to improve quality. I will now provide a few examples from industrial applications.

Example 6. SKF, the multinational Swedish ball bearing manufacturing company, experienced some years ago problems with a particular bearing type, custom-made in large volume to a particular customer, see Hellstrand (1987). Apparently, a competitor was able to show reliability data indicating that their ball bearing would last longer than SKF's. The customer therefore put SKF on notice that if they did not improve their bearings, they would loose this large and lucrative contract. To deal with this problem SKF assembled a team of engineers and a statistician. They ran a three factor two-level factorial experiment with factors heat treatment, osculation (essentially the contact area between the ball and the grove) and cage design. Accelerated life tests were performed with ball bearings made according to each of the eight factorial combinations. When analyzing the data, the team found that because of a large interaction effect between heat treatment and osculation, previously not know to SKF, they were able to extend the bearing life by a factor of five. This spectacular discovery was later confirmed with follow-up experiments, documented and presented to the customer, and the contract for future delivery was maintained with SKF. This is a good example of the power of factorial experiments, and in particular their ability, in few trials, to provide information about complex cause-and-effect relationships such as interactions between product design factors and relevant quality characteristics. It is also a prime example of how to design quality in upstream.

Example 7. Robustness, making products insensitive to hard-to-control variation, is an important new idea promoted much by Dr. Genichi Taguchi from Japan. In developing robust products, factorial experiments play a crucial role. To give a simple example, a hearing aid manufacturer had for years worked on miniaturizing the size of their hearing aid. However, this caused problems with placing knobs for volume control and other functions large enough, especially for older customers, to comfortably operate the hearing aid. A remote control was therefore developed, but turned out to have some problems with reception. It was conjectured that it possibly had to do with which cover was used, their own or one supplied by a vendor. Another conjecture was that the receptor coil in the hearing aid varied in production and therefore did not pick up the signal send at a single frequency of 24 kHz. However, this variation would be very difficult to control in production. It would be better if the system could be made robust to this variation. An alternative but slightly more expensive transmission program send out a signal sweep ing through a frequency range from 24 kHz to 25 kHz. A factorial experiment was run and revealed a large interaction effect between the transmission program (a steady 24 kHz signal versus a sweep signal between 24 to 25 kHz) and the controlled variation in the receptor coil. Moreover, the type of cover turned out to have no effect. This meant setting the easy-to-control factor, the transmission program, such that the influence of the hard-to-control factor, the receptor coil variation, was minimal. Thus the product was made more robust. This is, as indicated above, a very simple example of robust product development. However, the idea has significant implications. Quality engineers are no longer relegated to after-the-fact down stream inspection assignments, but now take a prominent seat around the table when new products are being developed. Their charter now is to assure that quality is designed into products at the design stage. This often requires a large experimental effort, where the expertise of quality engineers is crucial. The following example will further illustrate this idea.

Example 8. A large paper company making consumer products had a few years ago run into serious problems getting a new machine to run well enough to produce a new product, because of excess variation. By the time they asked for help from a statistical consultant they were 8 months behind schedule and had spend an estimated $500,000 desperately trouble shooting the process. After a series of seven or eight factorial experiments had helped remove a large portion of the variation, it became clear that the remaining source was an unexplainable but now very visible cyclic pattern of variation. After some further experiments, some of which essentially only kept the experimenters purposefully occupied, they somewhat serendipitously discovered that the cycle frequency could be changed by changing the length of the conveyer belt. After a few more experiments, this led them to the conjecture that the cycle was created by the length wise flexing of the conveyer belt. When the belt material was changed to one much less flexible, the cycle disappeared, the products were acceptable, and as an added benefit, it was possible to run the machine 50% faster than originally thought without adversely affecting quality. This was an important discovery because the machine was the prototype for an originally estimated total of seven machines. With this higher output per machine fewer were necessary and hence the company saved several million dollars in investments in new equipment. As it turned out, the company had many other similar machines that to a less severe degree suffered from the same problem. When they applied the same experimental strategy to these, they were able to get more consistent quality and a higher output with the obvious economic benefits as a result.

WIDER APPLICATION OF EXPERIMENTATION

The experiments that I have talked about so far have been applied to manufacturing processes or product development. However, experimentation is also applicable to service. Unfortunately, we do not at the present time have many well-documented examples available. However, I have been told by Dr. George Box about a few executed, as I understand it, with help from the consulting company QualPro. These examples might stimulate others to apply similar ideas. For example, credit card institutions in United States regularly send out large bulks of letters to get people to sign up for their cards. To increase the response rate, a particular credit card company first hired a consulting company, experts in consumer behavior, to help out. They suggested various ways to improve the response rate, that later appeared to have had little if any effect. The company then decided to use a large scale factorial experiment where factors like the color, size of the envelope, the text and various other factors were systematically changed. This lead to the discovery of a certain factorial combination that increased the response rate by a few percentage points. That might not sound like much, but in the credit card business, a one percent increase in new credit card subscribers might on a nation wide basis translate into millions of dollars in new business. Another example was an insurance company that experimented with various policies in the settlement of claims. Traditionally, insurance companies will make it difficult for policy holder to receive compensation for injuries. Moreover, they try to get the claimants to sign documents, once they settle, that no further claims can be filed should other injuries surface later. Such policies naturally build up animosity towards the insurance company and encourages law suits. However, from a factorial experiment conducted over a period of time using different combinations of settlement policies, an insurance company found that if they treated the claimants very generously up front, the company greatly minimized the number of extremely costly law suits filed later by disgruntled customers.

Let us now take this one step further. Systematic experimentation need not be confined to the private sector. There is a growing interest in using TQM concepts in the public sector and to Community Quality Improvement. Why not use scientific approaches to test the effect of social policies, administrative procedures, police work, pollution reduction strategies, educational programs and how to improve local communities? We might find out that there are much better and cheaper ways of getting people back to work, helping the elderly, reducing crime, providing health services and preventive medicine.

LOOKING AHEAD

Quality improvement is about learning — learning faster and more systematically. Deming's PDCA cycle is about systematic learning. Knowledge comes from studying a system, and with knowledge we can predict and control what will happen. But what is more important, we can, with insight, understand how to change systems so they better serve the customers. The insight we get from statistical studies and experiments will likely enable us to generate new ideas about how to create better systems that we would not otherwise have thought about.

The role of statistics is to help identify patterns and provide information through observation, surveys, and experiments about the needs of customers, how to design products or services that will satisfy those needs, for planning production, in managing operations, in selling products or services, in after-sales service, and in service organizations in tailoring services to the market.

Unfortunately, traditional teaching of statistics encourages a too mechanical application. The focus is often on mathematical details. Students learn, often from uninspired teachers, how to plug numbers into formulae and not the scientific process of discovery as indicated in the examples above. Too much emphasis is on testing hypothesis where it is often in practice much more important to generate the hypothesis in the first place and to get ideas about how we can deal with the system causes and not just fix special causes. As Deming and Juran claims, 80% of the problems are systems causes — obviously those are the ones we should work on.

Computers have already, and will continue to have, an immense impact on the use of statistics for quality improvement. It used to be that quality control applications were simply control charts and acceptance sampling. Today the whole toolbox of statistics is being used, or should be used — experimental design, response surface methods, multivariate statistics, sample surveys, reliability and life testing, statistical graphics and exploratory data analysis. Data are going to be more accessible. Automated data logging is already possible on a grand scale. Bar coding opens up immense possibilities for tracking quality problems. More things can be measured with increasingly sophisticated instruments. Larger data sets can be stored and scrutinized. Software packages will make it much easier to perform analysis. Patterns and assignable causes can be found. Correlation's can be discovered. New computer intensive styles of data analysis are possible. Several analyses are possible. Experimental data can be plotted in many more ways looking for patterns in residuals. Relations between cost and quality can be found. Economic and accounting data bases can be merged together with technical data bases to look for patterns of waste and cost of quality, and this may eventually help develop a better understanding of the economics of quality.

All of this is really not so futuristic as it might seem. The technology is here, and the data needed for scientific quality improvement is most often already available. What might be missing is the vision, the idea of initiating investigations and the statistical know how. What worries me about the direction of the Quality Movement at the moment is the increasing emphasis on rigid compliance with standards. Standards obviously play an important role. Whenever we find a new and better way to do something, we ought to institute it as a standard. However, as standardization is pursued in many companies that I have been in contact with, there appears to be a serious danger of forgetting that quality is ultimately about satisfying the customers and not just the auditors. Hopefully this article will stimulate ideas about how the quality movement can return to a healthy balance where customers are in focus and scientific method is the catalyst for achieving continuous never-ending learning and improvement.

ACKNOWLEDGMENTS

Although the opinions expressed in this manuscripts are mine, I want to acknowledge the importance of many discussions with Dr. George Box in helping me shape the views presented in this paper. Dr. Spencer Graves has also served for years as a helpful sounding board. This work was sponsored by the National Science Foundation grant number DMI 9500140.

REFERENCES

Beveridge, W. I. B. (1980), *Seeds of Discovery*, New York: W. W. Norton and Company.

Bisgaard, S. (1989), "The Quality Detective", *Philosophical Transactions of the Royal Society of London.* Vol. A 327, pp. 499-511.

Box, G. E. P. (1976), Science and Statistics, *Journal of the American Statistical Association*, Vol. 71, pp. 791-799.

Bronowski, J. and Mazlish, B. (1960), *The Western Intel lectual Tradition*, London: Hutchinson of London.

Butterfield, H. (1962), *The Origins of Modern Science*, New York: Collier Books.

Crouch, T. (1989), *The Bishop's Boys: A Life of Wilbur and Orville Wright,* New York: W. W. Norton & Company.

Deming, W. E. (1986), *Out of the Crisis*, Center for Ad vanced Engineering Study, MIT.

Dreyer, J. L. E. (1963), *Tycho Brahe*, Reprinted by Dover Publications, Inc. New York.

Hellstrand, C. (1989), "The Necessity of Modern Quality Improvement and some Experience with its Implementation in the Manufacture of Rolling Bearings", *Philosophical Transactions of the Royal Society of London*. Vol. A 327, pp. 529-537.

Ishikawa, K. (1985), *What is Total Quality Control? The Japanese Way,* Englewood Cliffs, NJ: Prentice-Hall.

Juran, J. M. (1992), *Juran on Quality by Design,* New York: The Free Press.

Shewhart, W. A. (1931), *Economic Control of Quality of Manufactured Product*, Princeton: D. Van Nostrand Company, Inc.

Shewhart, W. A. (1939), Statistical Methods From the Viewpoint of Quality Control, Edited by W. E. Deming, reprinted by Dover Publications, Inc. New York.

Tames, R. (1979), *Josiah Wedgwood*, Aylesbury, Bucks, UK: Shire Publications.