Lessons from a Career in Quality

Edward G. Schilling

How great it is to come back to Knoxville. Some years ago, while at GE in Cleveland, I was on the Advisory Board for the statistical program at the University of Tennessee. It was just starting its emphasis on SPC, and what wonderful progress has been made since then. Dave Chambers would be proud indeed to know how the field has evolved, and how UT has evolved with it.

When asked to speak, I wondered what to speak about. It was suggested that since this is for long service in the field, I should talk about the development of industrial statistics, and about some of my experiences along the way. As I thought about this, I did the math and realized that this year is exactly my 50th anniversary in the field, and so that choice demonstrates more insight than we could have imagined. So let's talk about that 50 year period and some of the lessons that came out of it. Most of which demonstrate the importance of non-numerical aspects of statistical applications and especially communication.

Speaking of communications, I recall an article in the Toronto Globe and Mail some years ago about the American Statistical Association Convention, which was being held there. The writer reported that as he walked through the lobby of the headquarters hotel, he spotted two statisticians talking to each other. He went about his business and when he came back through the lobby, they were still there but they were both fast asleep. I tell you this not as a warning, but as an invitation to do whatever comes naturally; just don't fall off your seats doing it. It's been a long day, so relax and let's see what insight we can gain by looking over those 50 years and beyond.

Now, in terms of the development of the field in industrial statistics, it goes back further than I do, but I believe we can distinguish three major phases: Theoretical development (1920-1950), Implementation (1950-1980) and Popularization (1980—). Each of these phases can be characterized by the techniques involved. Theoretical Development involved the invention of new techniques: The control chart, various sampling plans, techniques of reliability, and most importantly, experiment design. It lasted from the seminal work of Fisher, Shewhart, Dodge and others to the short courses in SQC given in World War II. By the end of the war the foundation had been laid, but its use was most often confined largely to knowledgeable individuals and small groups within a relatively small number of companies, such as General Electric, Hamilton Standard, DuPont, Eastman Kodak and the like and it was done on a project by project basis.

After the war came a period of practical implementation lasting roughly until the 1980's. The work of Fisher in an agricultural setting was transposed by Box, Hunter, Daniel, and others into factorial and fractional factorial designs, response surfaces, EVOP etc. The work of Shewhart was supplemented by CUSUM by Page and EWMA by Roberts. Industrial Quality Control magazine, the backbone of the quality movement, was transformed into the Journal of Quality Technology and Technometrics set a new standard for articles in the field. We were asked "If Japan can, why can't we? The seeds had been sewn for an expansion of understanding and use of statistics around the world. Then came the computer and the next phase was upon us; we are living in it and it is popularizing statistics.

What the computer revolution has done for us is to unburden the statistician from the chains of tedious calculation into a period where data manipulation and analysis is almost instantaneous. This frees the statistician to be what he or she should have always been, a philosopher and an expert in inductive logic rather than a mechanical robot enslaved by the complexity of the techniques involved, The philosophy of statistics is growing more and more useful and responsive in the real world. The TV networks are showing what amounts to confidence intervals on their polling results. Sophisticated exit polls have begun to be so well regarded that they are alleged to alter elections, so much so that the results are being withheld until the polls close. Statistics is becoming imbedded in other disciplines. Statistical thinking is becoming the norm, and is being taught at the grade school level.

This is the period of popularization of statistics. How exciting! The statistician is no longer thought of as an associate of the broadcaster at a football game. We no longer have to explain what it is that we do. It is no longer necessary to use the prefix "mathematical" before the term statistics. The doors are open to a blossoming of statistical methodology in industry and the outside world as well. Indeed, standards activities are helping to popularize and standardize the discipline. International standards already exist in the areas of confidence intervals, Design of Experiments, SPC, Acceptance Sampling, Measurement Systems, Interlaboratory Testing, etc. − and there even is a technical report on how to take a random sample.

I looked at those phases and my 50 years, and came to realize that in living through them, I had witnessed first-hand the development of industrial statistics from the inside and from the bottom up. I had seen it from a variety of vantage points, from a quality technician at a semiconductor plant, to a division quality manager with 43 plants to deal with, and from the academic, as well as the industrial side. I had watched the technology of statistics change from highly personal individual analysis with pencil and paper, to an output on computer paper, the kind with holes on the side and later to 8 1/2 x 11 sheets. The mechanics of analysis changed also from a hand crank mechanical device to the hand held calculator, to computers imbedded with even the most complicated of statistical techniques.

Few recognize names long since passed by like Monroe, Marchant and Frieden (they were desk calculators which allowed the user to add, subtract, multiply and divide, but not take square roots); great in the grocery store and of real value to the statistician who had nothing else to rely on. And who remembers the people enshrined in the names of techniques or awards, or books that are found only in the dusty part of the library. These were real people who gave so much to the field. I was lucky to be able to know and work with some of the early leaders in industrial statistics, such as Ott, Dodge, Wescott, Juran, Bicking and others, and it was a pleasure to do so. There were lessons to be learned. Let me illustrate with some of my experiences along the way. These are not profound, but may be instructive. But, most importantly, they are all true stories – I'll vouch for that.

My introduction to statistics was in the army those fifty years ago. I was a member of a group of enlisted men which was directed to use "scientific methods" to establish the number of people necessary to staff various units. We all had master's degrees or above. I had been a time study man and held a pilot's license, so I was assigned to fixed wing aircraft maintenance. These were used as spotters for the artillery among other duties. We took time studies, checked records and did our best to determine the staffing requirements.

One day the colonel who was in charge announced that he had sent an "all points" bulletin requesting old maintenance records from army units around the world. Now we were working in a very large room, about the size of a basketball court. The records came in and were filed. The files advanced around the walls of the room until no room was left, and they had to be sent to storage. Obviously, with hindsight, we saw that we had provided an outlet for units to discard their old maintenance records and they did – in abundance.

Now what has that to do with statistics? We couldn't deal with that volume of data (the computer was not with us then) and so we had to sample. After some nights in the library brushing up on the few statistics courses we had, we went to work taking random samples and calculating averages, standard deviations, and other measures. We thought we were sampling the population of army fixed wing aircraft, but I realized later that we were sampling from a frame consisting of those records that had been sent in from people who wanted to get rid of them, which may or may not have been representative of the true population of aircraft around the world. If only we had had Dr. Deming; he would have spotted the discrepancy immediately. Fortunately, it did not affect the staffing to any large degree. That is how I became interested in statistics, and learned the importance of how we collect the data and that the relation of the sample to the population is a vital consideration in any statistical analysis.

Some of the most helpful lessons I learned were while I was a Quality Technician at a semiconductor plant in New Jersey. Ours was a process control group. (Yes, they did have process control back in 1960). I was assigned to a very astute quality engineer, Carl Mentch, who taught me never to fight battles I knew I would lose, but never give up – there is always another day. Carl was a student of General Gebhard von Blucher, who lost seven straight battles to Napoleon before running up Napoleon's right wing at Waterloo, sealing the allied victory. He also insisted that we cost out all our projects. He understood the significance of profit in an industrial enterprise, a lesson which helped me later as a quality engineer.

What has this to do with statistics? Everything! Acceptance of our ideas does not often come easily, and the bottom line contribution of our results is the key to convincing others. Statistical analysis in industry is worthless unless it gets action, and sometimes the action required does not come on the first try.

These were not the only lessons I learned from the engineers. For example, the variability of the mean of certain electrical parameters was to be held as closely as possible for the Minuteman missile. The rocket scientists (yes, there are such things) did not care about the level, which could be adjusted for as long as the standard deviation remained small and stable. But the process was not always in a state of control, so it was decided that all acceptable units were mixed with a large supply of previous units to create a more stable population. This population was then sampled for acceptance of lots of a much smaller size taken from it. In effect then, the quality engineer built a super-population with a fairly stable variance which could then be sampled. The lesson here is to be open minded and creative in your applications.

This was also where I learned the value of the control chart for maintaining control over an industrial process. In those days, they used mechanical gauges for micro measurement of the thickness of semiconductor materials at a certain stage of manufacture. Suddenly the $\overline{X}$ chart went out of control, while the R chart remained in control. Weeks went by; the cause could not be isolated. Suddenly the problem was solved! An astute operator plotting the charts determined the cause. The gauges were set up near a door. On occasion the door remained open, allowing a gentle breeze to waft over the gauges, and these very sensitive devices exhibited measurement error. The lesson here is that the beauty of the control chart is its simplicity so that non-statisticians are at ease using it, and are willing to share the results.

Incidentally, as a student working in a foundry I was stationed at a conveyer belt to sort pieces which came from the shake-out of molds farther up the line. I was immediately behind an older gentleman who was straightening rods on an ancient but serviceable machine; however there were breakdowns of his machine as many as two or three times a day. Maintenance would fix it in a half hour or so while the gentleman watched or went away for a while. Then I realized that the breakdowns of the machine came whenever the operator wanted to take a break himself. Obviously he knew more than Maintenance about his machine and put his knowledge to use. The lesson here is never to underestimate the knowledge or shrewdness of an operator.

Probably the most valuable lesson also came from my experience as a quality engineer. You learn a lot when you are there on the spot to gather the data, do the calculations, disseminate them as appropriate, and try to make sense of the results. Now the engineer in the next operation was up for promotion to Quality Manager of our group. To show that he was really doing a fine job, he decided to post his control charts on a bulletin board of some size. It had to be big since he had so many charts. The units of product were, at that time, actually sold from one production area to the next. Sampling plans, usually attributes plans, were used to insure that the product was acceptable before it was passed on.

His $\overline{X}$, R charts for various electrical parameters were set up using the sampling plan data to avoid re-inspection and they were almost all out of control. Weeks passed, and he was unable to stabilize the charts, so much so that another engineer was promoted to Quality Manager instead. After a few months more of instability, he left the company. After taking a few more courses for my Masters Degree at Rutgers, I realized what had happened. By using the attribute sample sizes of 20 or 50 or 80 to plot $\overline{X}$, R charts, the charts became so sensitive that they were picking up the effect of assignable courses too small to be identified. It was then that I appreciated Shewhart's work in going to the Hawthorne plant to determine a practical sample size for the control chart, namely 4 to 9 – a long way from 20, 50,

80. That experience illustrates the importance of the statistical power curve of sampling plans, of control charts and of tests of hypothesis as well. It can be just as bad to have samples too large for their purpose as to have samples too small to distinguish important differences.

After obtaining my masters degree from Rutgers, I went to work as a Senior Quality Engineer at the Carborundum Company, maker of abrasives of various kinds and shapes. This requires a lot of sifting and screening and a considerable amount of mixing. I learned disdain for mixing, which requires much time and care to do it right. We talk of drawing random numbers from a hat, but not the complexity of how to do the drawing. I am reminded of the drawing held in Congress for the first individuals to be drafted for the Korean War. What a great opportunity for the politicians, and as you might guess, the proceedings were televised. The names were put in a bowl and the results of the drawing were published for all to see. You can be proud as individuals interested in statistics, for when the data was analyzed the next day, the drawing was shown to be non-random indeed. Why? Because whoever put the names in the bowl did so, as you rightly expect, in alphabetic order and the mixing that was done was totally inadequate, so beware of mixing, and when you do it, do it right.

Speaking of non-random events, it was also at Carborundum that I learned the power of outliers to affect our analysis. A plant was closing down. It had produced zirconium and hafnium. The vice president for research asked us to take the archived test data for various chemical tests and to build a model so that if the plant was reopened, the model could be used for the setup of the plant. This suggested multiple regression, and so we blindly went ahead. The results were disappointing to say the least, with R^2 at the bottom of the scale. When we looked at the scattergrams we could see why. Outliers were all over the place. We continued the study with limited results. When doing regression, always look at the scattergrams first and be careful of fishing projects such as this, and always beware of outliers in whatever you do.

I worked in a group where the rapport with the plants was excellent, that is until one of the statisticians was made a "persona non grata" at one of the plants. This was the result of a supercilious attitude, which bordered on arrogance. While this is extreme, too often the statistician talks above the heads of the operating personnel, which results in a no-win relationship in which the statistical results are lost or ignored. We must treat those with limited statistical knowledge with respect, and talk their language. Harold Dodge asked me to deliver his Grant Award speech at the ASQ Convention in Cleveland, since he was unable to make the trip. The title of his speech was "Keep it Simple" − good advice for any statistician.

Some years later, while a professor at Rochester Institute of Technology, I did some consulting work with a petro-chemical company which wanted to become certified to ISO 9001. We did an audit of the quality system in one of the plants. As part of this activity, one day we walked into the control room for one of the processes. The operator was seated at a console which displayed several Shewhart control charts. "Do you use the control charts?" we asked. "Oh no" was the reply "They are always out of control, and we rarely find the cause". The charts gave an audible signal when they went out of control, but why were they constantly giving signals? We looked into the software and found that the program was applying not only the 3σ rule, but 5 other rules for out of control in addition (such as 7 points in a row on one side of the median, etc.). This increased the $\propto$ risk on the charts and led to the questionable signals. The extra rules were dropped, solving the problem. This is a case where investigating the quality system led to important statistical results, and another lesson in the importance of power in statistical tests.

Did you ever attend the morning meeting at a chemical company where they review the previous night's results? I did and it was very instructive. Everyone had their own theory about what happened. The process went up by a unit or down by three. There was no concept of common causes, let alone the standard error. It was the classic illustration of lack of appreciation of statistical thinking that people talk about. The beauty of statistical methods is that when they are brought to bear in a meeting such as this, everything stops and concentration is on the numbers, and not on recrimination and discourse. Statistical thinking is essential.

Consulting can teach you a lot: I did some work for a nationally known jewelry company out of New York City. They were situated on two floors high up in a skyscraper. The main floor held their office, while they made and repaired jewelry on the floor below. Now this was a national chain, so they had plenty of employees, but most jewelry manufacturers are mom and pop operations with say from four to eight people. They were doing 100% inspection, and I was asked by the Vice President for Quality to set up some sampling plans in an effort to be more cost effective. I always ask "When you reject the product, what do you do with it?" (A neat question to ask an ISO 9000 audit). "Oh, we send it to the floor below, repair it and sell it. "You don't send it back?" "No." "Do you notify the manufacturer of the defective product?" "No". So I asked that they notify the manufacturers of recent defective material. You guessed it, the managers of the defective material rang the phone off the wall with apologies since they had no idea their product was defective. There was a dramatic increase in the quality. The lesson learned here was one of feedback. Those responsible for action must be notified so that corrective measures can be taken.

There may not have been control charts if it were not for Shewhart's Bowl. Shewhart was a very practical man. He knew of the consequences of the central limit theorem with very large sample sizes as predicted by theory, but it was thought at that time that sample sizes had to be very large to produce the effect. He wondered what the effect would be with small sample sizes. Accordingly he made up sets of chips representing various populations, put them in a large brown bread mixing bowl, took samples, and developed the sampling distributions for relevant statistics. He found that the central limit theorem held surprisingly well for sample sizes even as low as four or five and even for some representative distributions with non-normal shapes. So by plotting the mean of four or more, we could use the normal distribution to develop probabilistic limits which would tell us when a process changed in level, and the control chart was born.

Ellis Ott was a good and respected friend of Dr. Shewhart and on his passing, Shewhart's wife gave the original brown mixing bowl to Ellis. When Ellis passed away, his wife Virginia asked me what to do with the bowl. I thought of the Smithsonian, but when I called to discuss the matter with the American Society of Quality, I found that Shewhart's secretary had given the chips to the Society, and they were in the ASQ's vault. It was clear that the bowl belonged with the chips. Ed Beasley, a colleague of mine at G.E. had a cottage near the Ott summer home in the Adirondacks. He would bring the bowl to Cleveland and I would then carry it between my knees by air to Milwaukee, and so it was that the Shewhart bowl and chips now reside in a permanent display at ASQ headquarters. The lesson here was that theoretical results must match physical reality, and if they don't, it is better to question the theory than reality.

My last illustration concerns integrity. I headed a divisional quality group which did statistical consulting as well as the usual quality activities such as quality costs, divisional reports, audits, etc. It was suggested by one of the executives that we use our statistical consulting to inform top management of the status of sensitive programs before the responsible department did. This amounts to ratting on the responsible department. This would have killed our consulting activity in no time, besides the ethical problems involved. We never did, and our statistical consulting continued to thrive. It is essential that other people's data be treated with respect, and that the statistician and the practitioner cooperate in an atmosphere of mutual trust and confidence.

We have looked at statistical analysis from both a micro and macro level. I believe there is more to statistics than just the numbers. A good analysis will get quick and effective action only if it is understood and accepted by non-statisticians as well. Fortunately, the six-sigma programs based on statistical thinking are bridging the gulf of understanding between the statistician and the operating people, but more has to be done. The need for professional statisticians to supplement the basic methods of six-sigma must be recognized. Universities like the University of Tennessee, Virginia Polytechnic and RIT, among others, are turning out the next generation of statisticians who will find themselves in an atmosphere more conducive to use of statistics than ever before. I believe 6σ will continue because it is organizationally driven rather than the helter-skelter operations that preceded it. We are entering the period of Popularization of Statistics, and the world will be the better for it.

There is great potential for statistics in the 21st century but, as you may guess, some people will miss the mark. There was not enough room for another statistics course in the RIT Engineering program, so I suggested that our Center team up with the Philosophy Department and give an inductive logic course with them giving an overview of inductive logic, and our Center for Quality and Applied Statistics providing the statistical training. It could then be slipped into the Engineering curriculum as an elective. I went to the Philosophy Department with the idea. Surely the philosophers would want to sponsor a high minded approach such as this. Their response – "Who will pay for it?"

These lessons show there is much more to statistical analysis than just the numbers. We have seen illustrated such non-numeric considerations as:

1) The relation of the sample to the population as a frame must be known
2) Cost considerations and the need for action is imperative
3) Open-mindedness and thinking "out of the box" is always a plus
4) Simplicity rules
5) Never underestimate the operator
6) The power of a test is vital
7) Insufficient randomization is deadly
8) Beware of outliers
9) Talk in the language of the user
10) Beware of compounding multiple tests
11) Statistical thinking is essential
12) Feedback of results is vital
13) Make sure theory matches reality
14) Treat other people's data with respect

Now go out and get some data, but not too much.
And always remember:

Statistics is not for the close-minded who would use the data as a means to an end,
But for the open-minded whose end is what this data means.

Thank you.